PROGRESS OF THEORETICAL PHYSICS

Founded by H. Yukawa in 1946

Volume 11

JANUARY—JUNE

1954

CONTENTS

No. 1   JANUARY

No. 2.   FEBRUARY

No. 3   MARCH

Nos. 4, 5   APRIL, MAY

No. 6   JUNE
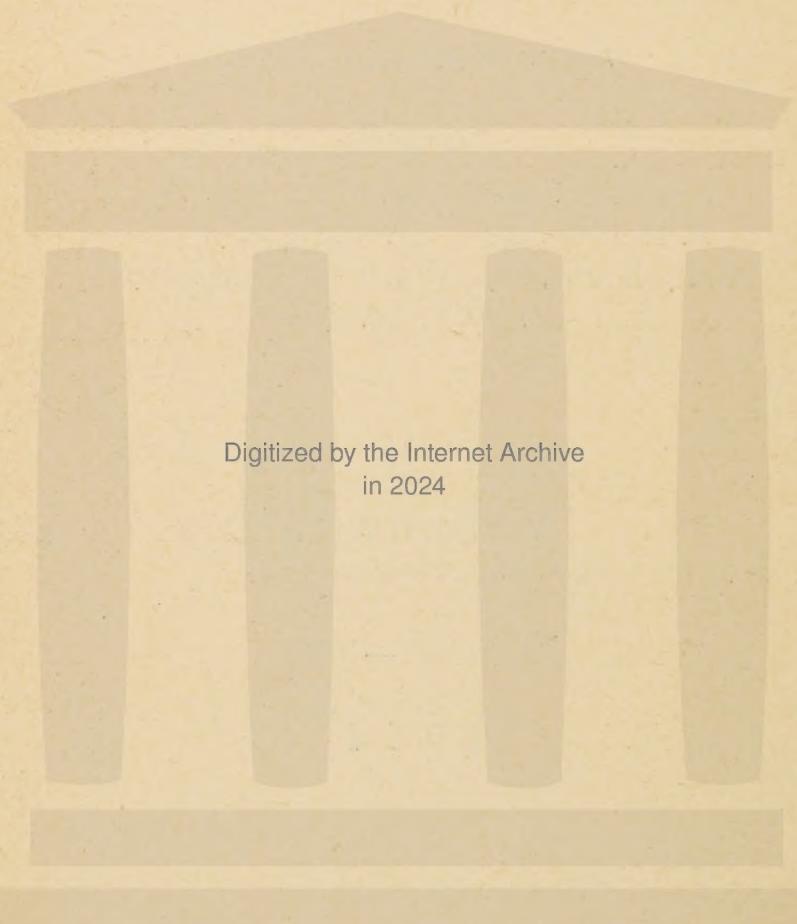
Invariant Delta Functions in the Sense of Distributions*

Takehito TAKAHASHI

Seizyo University, Setagaya, Tokyo

(Received November 16, 1953)

Invariant delta functions are most adequately interpreted in the sense of distributions of L. Schwartz. They are expressed as the sum of proper distributions and mass-dependent point functions. First terms are interpreted as the logarithmic or finite parts of the divergent integrals corresponding to the inverse square of the four-dimensional distance. Point function term of $\varDelta^{(1)}$ exhibits logarithmic singularities on the surface of the light cone, defining a finite value as a distribution.

§ 1. Introduction

To deal with the queer " functions " $\varDelta$, $\bar{\varDelta}$, $\varDelta^{(1)}$ etc., and to clear up the ambiguities in the calculations beyond the usual conventional technics, we must treat them throughout in a mathematically justified way.

The theory of distributions will be powerful to this purpose. So we are to interpret the invariant delta functions not as mere point functions but as a sort of distributions.

At first a rough sketch about the concept of distribution will be given. For details the original books of L. Schwartz[1] should be refered to.

A distribution on an open space V is a continuous linear functional $T(\varphi)$ defined for all infinitely differentiable complex valued function $\varphi(x)$, $x \in V$, whose carrier (the closure of the set of points $x \in V$ such that $\varphi(x) \doteqdot 0$) is any compact subspace of V. Here continuity is said with respect to all the derivatives of $\varphi(x)$. The symbol φ is used throughout this paper to denote the function with qualities mentioned above.

A partial derivative of a distribution T on R is defined by

$$(\partial/\partial x_i) T(\varphi) = -T(\partial\varphi/\partial x_i), \tag{1}$$

$$x = (x_1, \cdots, x_n) \in R^n.$$

Let S and T be two distributions. We define the convolution $S * T$ of S and T by

$$(S * T)(\varphi) = S_x(T_y(\varphi(x+y))), \tag{2}$$

when the right hand side makes sense.

Examples :

(i) Any function $f(x)$ $(x \in R^n)$, if summable on every compact set, defines a distribution T_f on R^n :

$$T_f(\varphi) = \int_{R^n} f(x)\varphi(x)\,dx.$$ (3)

In this meaning a point function itself may be looked upon as a distribution and we write $f(\varphi)$ instead of $T_f(\varphi)$.

(ii) Dirac's distribution ∂_{R^n}, defined by

$$\partial_{R^n}(\varphi) = \varphi(0).$$ (4)

Now assume that the function $F(t)$ defined for $t > 0$ (or $t < 0$) may be written in the form

$$F(t) = A + B \log|t| + \sum_{\Re(m) > 0} t^{-m}(A_m + B_m \log|t|) + E(t),$$ (5)

where the exponent m appear finite times and $\lim_{t \to 0} E(t) = 0$.

Then we name A and $-B$ as the finite part and the logarithmic part of $F(t)$ respectively and write

$$A = P_f F(t) \quad (\text{or } P_f F(0)),$$ (6)

$$-B = P_l F(t) \quad (\text{or } P_l F(0)).$$ (7)

For instance we can easily prove the following relations :

$$P_f \int_0^\infty \varphi(x) x^{-1}\,dx = \lim_{\varepsilon \to 0}\left(\int_\varepsilon^\infty \varphi(x) x^{-1}\,dx + \varphi(0)\log\varepsilon\right),$$ (8)

$$P_f \int_{-\infty}^0 \varphi(x) x^{-1}\,dx = \lim_{\varepsilon \to 0}\left(\int_{-\infty}^{-\varepsilon} \varphi(x) x^{-1}\,dx - \varphi(0)\log\varepsilon\right),$$ (9)

$$P_l \int_0^\infty \varphi(x) x^{-1}\,dx = \varphi(0),$$ (10)

$$P_l \int_{-\infty}^0 \varphi(x) x^{-1}\,dx = -\varphi(0).$$ (11)

Eqs. (8)–(11) define distributions $P_f^\pm x^{-1}$ and $P_l^\pm x^{-1}$ on R^1 by writing the left sides of these equations as

$$P_f^+ x^{-1}(\varphi), \quad P_f^- x^{-1}(\varphi), \quad P_l^+ x^{-1}(\varphi) \text{ and } P_l^- x^{-1}(\varphi)$$

respectively. (They are thought to be zero when the carrier of φ lies outside of the integral domain.)

We observe that the distribution $P x^{-1}$ or $\bar{P}_l x^{-1}$ defined by

$$P x^{-1} = P_f^+ x^{-1} + P_f^- x^{-1}, \qquad \bar{P}_l x^{-1} = P_l^+ x^{-1} - P_l^- x^{-1}$$ (12)

coinsides with the principal value of Cauchy or twice the Dirac's distribution ∂_{R^1}.

They are obtained by differentiating the functions :

$$P x^{-1} = (d/dx) \log|x|,$$ (14)

$$\bar{P}_l x^{-1} = (d/dx)\varepsilon(x),$$ (15)

where

$$\varepsilon(x) = sgn(x). \tag{16}$$

For the purpose of later use we give another expression which avails complex integral.

Let C^+ (or C^-) be the contour along the real axis in the positive (or negative) direction, being replaced under (or above) the poles by semi-circles of radius r, which will tend to zero.

Then we can observe

$$\int_{C^+} \varphi(z) z^{-1} dz = P x^{-1}(\varphi) + (\pi i / 2) \bar{P}_t x^{-1}(\varphi), \tag{17}$$

$$\int_{C^-} \varphi(z) z^{-1} dz = - P x^{-1}(\varphi) + (\pi i / 2) \bar{P}_t x^{-1}(\varphi), \tag{18}$$

left sides of which define distributions $P_C^+ z^{-1}(\varphi)$ and $P_C^- z^{-1}(\varphi)$ on the complex plane. Thus we obtain

$$P x^{-1} = (1/2) (P_C^+ z^{-1} - P_C^- z^{-1}), \tag{19}$$

$$\bar{P}_t x^{-1} = (1/\pi i) (P_C^+ z^{-1} + P_C^- z^{-1}). \tag{20}$$

Analoguous arguments are possible in the space of higher dimensions. We shall investigate the integral $\int \varphi \cdot \sigma^{-2} dR^4$ $(\sigma^2 = t^2 - x_1^2 - x_2^2 - x_3^2)$ in the interior and exterior of the light cone and show that its logarithmic or finite part corresponds to the proper distribution term of the invariant delta functions (§ 4). Explicit expressions of them, together with the point function terms, will be determined in a natural way from the classical solutions of the Cauchy's problems (§ 2, § 3). Because of this way of treating the problem, usual covariant expressions are avoided.

§ 2. Cauchy's problems and invariant delta functions

(i) *Cauchy's problem of the homogeneous equation*

Let $L[u]$ be a linear differential form with respect to the three-dimensional space coordinates; consider following two problems of Cauchy

A) $\partial^2 u / \partial t^2 - L[u] = 0$ $(t > 0)$, \hfill (21)

$u_{t=0} = 0,$ $(\partial u / \partial t)_{t=0} = f(x).$ \hfill (22)

B) $\partial^2 G^R / \partial t^2 - L[G^R] = 0$ $(t > 0)$, \hfill (23)

$G^R_{t=0} = 0,$ $(\partial G^R / \partial t)_{t=0} = \delta_{R^3} .$ \hfill (24)

As is easily verified by direct substitution, the solution of A) is expressed as the convolution on R^3 of the initial value of $\partial u / \partial t$ and the solution of B) :[1) 2)]

$$u(x, t) = f * G^R(x | t), \tag{25}$$

and defines a distribution

$$u(\varphi) = \int_{R^3} u(x, t)\varphi(x)dx = \int_{R^3} f(y)\, G_x^R(\varphi(x+y))\, dy\,, \tag{26}$$

where
$$x = (x_1, x_2, x_3), \qquad y = (y_1, y_2, y_3),$$
$$dx = dx_1 dx_2 dx_3, \qquad dy = dy_1 dy_2 dy_3\,.$$

We shall avail the eq. (26) to determine the distribution $G(\varphi)$ from the knowledge of the solution $u(x, t)$.

In case of the wave equation

$$\partial^2 u/\partial t^2 - L[u] = -\Box u = \partial^2 u/\partial t^2 - (\partial^2/\partial x_1^2 + \partial^2/\partial x_2^2 + \partial^2/\partial x_3^2)u\,, \tag{27}$$

the solution of A) is given by the famous formula of Poisson :[3]

$$u(x, t) = (t/4\pi)\int_{\alpha^2=1} f(x+\alpha t)\, d\omega\,, \tag{28}$$

where
$$\alpha = (\alpha_1, \alpha_2, \alpha_3), \qquad \alpha^2 = \alpha_1^2 + \alpha_2^2 + \alpha_3^2,$$

and $d\omega$ means the surface element of the unit sphere.

Then we get

$$u(\varphi) = (t/4\pi)\iint f(x+\alpha t)\varphi(x)\, d\omega\, dx$$

$$= (t/4\pi)\int_{R^3} f(y)\, dy \int_{\alpha^2=1} \varphi(y-\alpha t)\, d\omega\,,$$

from which we obtain, on comparing with the right side of (26) and writing D^R for G^R,

$$D_x^R(\varphi(x+y)) = (t/4\pi)\int_{\alpha^2=1} \varphi(y-\alpha t)\, d\omega\,.$$

On putting $y=0$ we can determine the expression of $D^R(\varphi)$

$$D^R(\varphi) = (t/4\pi)\int_{\alpha^2=1} \varphi(\alpha t)\, d\omega = (r/4\pi)\int_{x^2=t^2} \varphi(x)\, d\omega\,, \qquad (r = \sqrt{x^2})\,. \tag{29}$$

Thus the distribution D^R means to make spherical mean multiplied by the radius of the sphere.

Next we consider the following case

$$\partial^2 u/\partial t^2 - L[u] = -(\Box + \varkappa^2)u\,, \qquad G^R = \varDelta^R\,. \tag{30}$$

The solution of A) is given as follows :[3]

$$u(x, t) = \frac{t}{4\pi}\int_{\alpha^2=1} f(x+\alpha t)\, d\omega + \frac{\varkappa}{4\pi}\int_{t^2 \geq y^2} f(x-y)\, \frac{J_0'(\varkappa\sqrt{t^2-y^2})}{\sqrt{t^2-y^2}}\, dy\,. \tag{31}$$

The second term of (31) being the convolution between point functions, we get immediately

$$\Delta^R = D^R + (\varkappa/4\pi \, S^+) \, J_0'(\varkappa S^+), \tag{32}$$

where

$$S^+ = \begin{cases} \sqrt{t^2 - \overline{x^2}} & (t^2 \geq x^2, \quad t \geq 0), \\ 0 & (\text{otherwise}) \end{cases} . \tag{33}$$

Take care that so far D^R and J^R are considered on K^3; the time variable t appears here as a continuous parameter.

(ii) *Inhomogeneous problem and elementary solutions*

C) $$\partial^2 v / \partial t^2 - L[v] = f(x, t), \tag{34}$$

$$v(x, 0) = (\partial v / \partial t)(x, 0) = 0 . \tag{35}$$

The solution of C) is expressed by the solution u of the homogeneous equation (i) A), as follows :[3]

$$v(x, t) = \int_0^t u(x, \tau ; t - \tau) d\tau , \tag{36}$$

where τ means the parameter introduced in the initial value of $\partial u / \partial t$; we write $f(x, \tau)$ instead of $f(x)$.

Assume that $f(x, t)$ and $G(\varphi)$ are both zero on any subspace of $t < 0$, on substituting (25) we can interpret (36) defining four-dimensional convolution :

$$v(x, t) = f * G^R(x, t). \tag{37}$$

Putting (37) in (34), we obtain

$$\partial^2 v / \partial t^2 - L[v] = f * (\partial^2 G^R / \partial t^2 - L[G^R]) = f$$

from which follows

$$\partial^2 G^R / \partial t^2 - L[G^R] = \delta_{R^4} . \tag{38}$$

Thus the distribution G^R (on R^4) proves to be the elementary solution of the operator $(\partial^2 / \partial t^2 - L)$.

§ 3.   Extension to the whole space and the distribution $\Delta^{(1)}$

In the usual manner we shall extend our theory in the domain $t < 0$. Consider the elementary solution G^A with carrier in the domain $t \leq 0$;

$$\partial^2 G^A / \partial t^2 - L[G^A] = \delta_{R^4} , \tag{39}$$

$$(G^A)_{t \geq 0} = 0 , \qquad (\partial G^A / \partial t)_{t=0} = -\delta_{R^3} . \tag{40}$$

If we set

$$G = G_A - G_R , \qquad \overline{G} = (G_A + G_R)/2 , \tag{41}$$

we observe that G and $\overline{G}$ satisfy following relations :

$$\partial^2 G/\partial t^2 - L[G] = 0 , \qquad (G)_{t=0} = 0 , \qquad (\partial G/\partial t)_{t=0} = -\delta_{R^3} , \qquad (42)$$

$$\partial^2 \bar{G}/\partial t^2 - L[\bar{G}] = \delta_{R^1} , \qquad (\bar{G})_{t=0} = (\partial \bar{G}/\partial t)_{t=0} = 0 . \qquad (43)$$

Explicit expressions of Δ^A, $\bar{\Delta}$, Δ etc. may be easily obtained.

Now let $P\tau^{-1}$ be the distribution defined on any straight line with time-like direction, then we define $\pi\Delta^{(1)}$ as the convolution of Δ and $P\tau^{-1}$:

$$\Delta^{(1)} = (1/\pi)\Delta * P\tau^{-1} . \qquad (44)$$

On substituting the expression of Δ :

$$\Delta = D + (1/4\pi) W , \qquad (45)$$

$$W = -\varepsilon(t) x S^{-1} J_0'(xS) , \qquad (46)$$

where

$$S = \begin{cases} \sqrt{t^2 - x^2} & (t^2 > x^2) , \\ 0 & (t^2 \leq x^2) , \end{cases} \qquad (47)$$

and taking the direction of τ on the time axis, we obtain

$$\Delta^{(1)} = D^{(1)} + (1/4\pi^2) P \int_{-\infty}^{\infty} W(\tau)(t-\tau)^{-1} d\tau , \qquad (48)$$

where $D^{(1)}$ is defined by

$$\pi D^{(1)}(\varphi) = (D * P\tau^{-1})(\varphi) = D\left(\int_{-\infty}^{\infty} \varphi(x, \tau)(\tau - t)^{-1} d\tau\right) \qquad (49)$$

and will be treated in the next section.

The second term of $\Delta^{(1)}$ can be calculated as follows :

$$P \int_{-\infty}^{\infty} \frac{W(\tau)}{t-\tau} d\tau = -P \int_0^{\infty} \frac{x}{s} \frac{J_0'(xs)}{t-\tau} d\tau - P \int_0^{-\infty} \frac{x}{s} \frac{J_0'(xs)}{t-\tau} d\tau$$

$$= P \int_0^{\infty} \frac{x}{s} J_0'(xs) \left\{ \frac{1}{\tau - t} + \frac{1}{\tau + t} \right\}$$

$$= P \int_0^{\infty} \frac{x}{s} J_0'(xs) \frac{2\tau \, d\tau}{\tau^2 - t^2} . \qquad (50)$$

Set

$$\sigma = (\tau^2 - x^2)^{1/2} , \qquad \sigma_0 = (t^2 - x^2)^{1/2} , \qquad (51)$$

then (50) becomes

$$P \int_{-8}^{\infty} \frac{W(\tau)}{t-\tau} d\tau = P \int_0^{\infty} \frac{x}{\sigma} J_0'(x\sigma) \frac{d\sigma^2}{\sigma^2 - \sigma_0^2} . \qquad (52)$$

To calculate this integral we set

$$x\sigma = \xi , \qquad x^2 \sigma_0^2 = \begin{cases} a^2 & (\sigma_0^2 > 0), \\ -a^2 & (\sigma_0^2 < 0). \end{cases} \qquad (53)$$

Then we have to consider the integral I_+ or I_- defined by

$$I_+ = P \int_0^\infty J_0(\xi)\,(\xi^2 - a^2)^{-1}\,d\xi, \tag{54}$$

$$I_- = \int_0^\infty J_0'(\xi)\,(\xi^2 + a^2)^{-1}\,d\xi, \tag{55}$$

according as the point in question lies in the interior or exterior of the light cone. The integrals $I_\pm$ can be evaluated by the aid of following contour integrals

$$\int_C z^2 (z^2 \mp a^2)^{-1} H_1^{(1)}(z)\,dz, \tag{56}$$

where C is the contour in the half plane $\Im(z) \geq 0$, consisting of semicircles $|z| = R$, $|z \pm a| = r$ and the parts of the real axis; second circles are not necessary in case of I_-. On letting $R \to \infty$, $r \to 0$, the integral along the real axis becomes

$$2(\mp a^2 I_\pm + 1).$$

Then we obtain, after calculating the residues,

$$I_+ = (\pi/2a)\,N_1(a) + (1/a^2), \tag{57}$$

$$I_- = (1/a)\,K_1(a) - (1/a^2), \tag{58}$$

where N_1 is the Neumann's function and $(-2/\pi)K_1$ is the Hankel's function with imaginary argument.

Thus we have completed the calculations of the second term of $\varDelta^{(1)}$, and we get

$$\varDelta^{(1)} = D^{(1)} + \begin{cases} (x/4\pi\sigma)\,N_1(x\sigma) + (1/2\pi^2\sigma^2) & (\sigma^2 > 0), \\ (x/2\pi^2(-\sigma^2)^{1/2})\,K_1(x(-\sigma^2)^{1/2}) - (1/2\pi^2\sigma^2) & (\sigma^2 < 0). \end{cases} \tag{59}$$

Notice that the second term, being a point function, exhibits logarithmic singularities on the surface of the light cone.

§ 4. Invariant delta functions as the logarithmic or finite part of σ^{-2}

We shall investigate the divergent integral

$$\int \varphi(x, t)\,\sigma^{-2}\,dR^4$$

with singularities on the surface of the light cone.

Let V_ε^+ be the domain defined by

$$(1-\varepsilon)^2 t^2 - r^2 \geq 0, \qquad 1 > \varepsilon > 0. \tag{60}$$

Then the integral for a fixed ε

$$\int_{V_\varepsilon^+} \varphi(x, t)\,\sigma^{-2}\,dR^4 \tag{61}$$

exists in the limit; first take $|t| \geq \delta > 0$ then let $\delta \to 0$.

The logarithmic part of the integral (61) defines a distribution $P_l s^{-2}$ and $\bar{P}_l s^{-2}$, as

$$P_l s^{-2} = (P_l s^{-2})_{t>0} + (P_l s^{-2})_{t<0} , \tag{62}$$

$$\bar{P}_l s^{-2} = (P_l s^{-2})_{t>0} - (P_l s^{-2})_{t<0} , \tag{63}$$

$$(P_l s^{-2})_{t \gtrless 0}(\varphi) = P_l \int_{V_\varepsilon^+, \, t \gtrless 0} \varphi(x, t) \sigma^{-2} dR^4 , \tag{64}$$

In the last expression $\varphi(x, t) = \varphi(ru, t)$ can be replaced by $\varphi(|t|u, t)$, if we observe that the integral

$$\int_{V_\varepsilon^+, \, t \gtrless 0} \{\varphi(ru, t) - \varphi(|t|u, t)\} \sigma^{-2} dR^4$$

remains finite when ε tends to zero.

Then we can calculate as follows:

$$(P_l s^{-2})_{t \gtrless 0}(\varphi) = \pm P_l \int_0^{\pm\infty} dt \int_{\alpha^2=1} \varphi(|t|u, t) d\omega \int_0^{|t|(1-\varepsilon)} r^2 (t^2 - r^2)^{-1} dr$$

$$= \pm P_l \int_0^{\pm\infty} dt \int_{\alpha^2=1} \varphi(|t|u, t) d\omega \left[-\frac{|t|}{2} (\log \varepsilon - \log(2-\varepsilon)) \right]$$

$$= \pm \frac{1}{2} \int_0^{\pm\infty} |t| dt \int_{\alpha^2=1} \varphi(|t|u, t) d\omega . \tag{65}$$

From eq. (65) we observe that

$$(P_l s^{-2})_{t<0}(\varphi(x, t)) = (P_l s^{-2})_{t>0}(\varphi(x, -t)), \tag{66}$$

and following relations hold:

$$D^R = (1/2\pi)(P_l s^{-2})_{t>0} , \qquad D^A = (1/2\pi)(P_l s^{-2})_{t<0} , \tag{67}$$

$$D = (-1/2\pi)\bar{P}_l s^{-2} , \tag{68}$$

$$\bar{D} = (1/4\pi)P_l s^{-2} . \tag{69}$$

Eq. (65) shows that the integral (61) has only logarithmic divergency, hence we can define the finite part of (61), by

$$(P_f^+ \sigma^{-2})(\varphi) = \lim_{\varepsilon \to 0} \left(\int_{V_\varepsilon^+} \varphi(x, t) \sigma^{-2} dR^4 + (P_l \sigma^{-2})(\varphi) \log \varepsilon \right). \tag{70}$$

On the domain V_ε^-:

$$t^2 - (1-\varepsilon)^2 r^2 \leq 0 , \tag{71}$$

we can evaluate the logarithmic part as follows:

$$P_l \int_{V_\varepsilon^-} \varphi \cdot \sigma^{-2} dR^4 = P_l \int_0^\infty r^2 dr \int_{\alpha^2=1} \varphi(ru, \pm r) d\omega \int_{-r(1-\varepsilon)}^{r(1-\varepsilon)} (t^2 - r^2)^{-2} dt$$

$$= P_l \int_0^\infty r^2 dr \left[\int_{\alpha^2=1} \varphi(ru, r) d\omega \int_0^{r(1-\varepsilon)} (t^2 - r^2)^{-2} d, \right.$$

$$+ \int_{\alpha^2=1} \varphi(r u, -r)\, d\omega \int_{-r(1-\varepsilon)}^{0} (t^2 - r^2)^{-2}\, dt \Bigg]$$

$$= - P_t s^{-2}(\varphi)$$

on employing the relation

$$\int_{0}^{r(1-\varepsilon)} r^2 (t^2 - r^3)^{-1}\, dt = (-r/2)(\log \varepsilon - \log(2-\varepsilon) + 2 - 2\varepsilon).$$

As before we can define the finite part

$$(P_f^{-} \sigma^{-2})(\varphi) = \lim_{\varepsilon \to 0}\left(\int_{V_\varepsilon^{-}} \varphi \cdot \sigma^{-2}\, dR^4 - P_t s^{-2}(\varphi) \log \varepsilon \right). \tag{72}$$

From eqs. (70) and (72) the distribution $P\sigma^{-2}$ is introduced by

$$(P\sigma^{-2})(\varphi) = (P_f^{+} \sigma^{-2} + P_f^{-} \sigma^{-2})(\varphi) = \lim_{\varepsilon \to 0}\left(\int_{V_\varepsilon^{+}} + \int_{V_\varepsilon^{-}} \right) \varphi \cdot \sigma^{-2}\, dR^4, \tag{73}$$

or explicitly

$$(P\sigma^{-2})(\varphi) = \lim_{\varepsilon \to 0} \int_{0}^{\infty} r^2\, dr \left[\int_{-(1-\varepsilon)r}^{(1-\varepsilon)r} + \int_{-\infty}^{-r/(1-\varepsilon)} + \int_{r/(1-\varepsilon)}^{\infty} \right] dt \int_{\alpha^2=1} \varphi \cdot \sigma^{-2}\, d\omega$$

$$= \int_{0}^{\infty} r^2\, dr \left[P \int_{-\infty}^{\infty} \sigma^{-2}\, dt \int_{\alpha^2=1} \varphi \cdot d\omega \right]$$

$$= \frac{1}{2} \int_{0}^{\infty} r\, dr \left[P \int_{-\infty}^{\infty} (t-r)^{-1}\, dt - P \int_{-\infty}^{\infty} (t+r)^{-1}\, dt \right] \int_{\alpha^2=1} \varphi(r u, t)\, d\omega, \tag{74}$$

which means, on writing t or τ for r or t,

$$(P\sigma^{-2})(\varphi) = \{(P_t s^{-2})_{t>0} - (P_t s^{-2})_{t<0}\}\left(P \int_{-\infty}^{\infty} \varphi(x, \tau)(\tau - t)^{-1}\, d\tau \right)$$

$$= -[(P_t s^{-2}) * P t^{-1}](\varphi). \tag{75}$$

Then we can deduce the relation

$$D^{(1)} = (-1/2\pi^2) P\sigma^{-2} \tag{76}$$

from eqs. (49), (68) and (75).

Next we shall show the relation

$$P\sigma^{-2} = -\square \log |\sigma|,$$

the right side of which defining the distribution

$$-\int_{R^4} \log |\sigma| \cdot \square \varphi \cdot dR^4. \tag{78}$$

On making use of the four-dimensional Green's formula on V_ε^{+} and the relation

$$\square \log |\sigma| = -\sigma^{-2} \qquad (\sigma^2 \neq 0),$$

the integral (78) is expressed as the sum of $P\sigma^{-2}(\varphi)$ (eq. (73)) and the surface integral

containing transversal derivatives. The latter cancels and we have the result.

Lastly we give the expressions corresponding to eqs. (19) and (20) of § 1 ; they are

$$P\sigma^{-2} = (1/2)(P_C^+ \sigma^{-2} - P_C^- \sigma^{-2}), \tag{79}$$

$$\bar{P}_l s^{-2} = (1/2\pi i)(P_C^+ \sigma^{-2} + P_C^- \sigma^{-2}), \tag{80}$$

where by $P_C^\pm \sigma^{-2}(\varphi)$ we mean the expression (74) whose time integral part is replaced by the contour integral on the complex t-plane along the contour $C^\pm$ defined in § 1.

These relations show that $(i/4\pi^2)P_C^\pm \sigma^{-2}$ are the distribution $D^\pm$ defined by

$$D^\pm = (1/2)(D \mp iD^{(1)}),$$

If the contour C' is altered by replacing the semicircle under the pole $t = -r$ by the semicircle above it, we obtain the contour C_F and $(-1/2\pi^2)P_C^F \sigma^{-2}$ defined as before coincides with $\bar{D}^F = D^{(1)} - 2iD$.

§ 5.   Conclusion

The problem of Fourier transforms[1] is not treated in this paper, but there is no difficulty to describe them as we have already succeeded to clear up the ambiguities in the expressions of delta functions.

At the present stage of quantum field theory we are confronted with many difficulties by treating proper distributions as point functions. Some of them may be got rid of if we can define delta functions first on any space-like surface then extend continuously to the whole space. Still it is meaningless, for instance, to argue about the values which delta functions will take at one space time point.

An interesting trial was made by W. Güttinger[1] to examine the calculations of the quantum field theory with distribution analysis, utilizing the arbitrary constants which appear by division process or by transforming the variable, but it seems to us that the whole field theory should be rewritten to fit the treatment so far made in the sense of distributions.

References

1) L. Schwartz, *Théorie des distributions*, I, II, Paris, Hermann (1950–51).
2) L. Schwartz, Ann. Inst. Fourier 2 (1950), 19.
3) R. Courant und D. Hilbert, *Methoden der mathematischen Physik*, II, Berlin, Springer (1937), 159, 165, 402, 448.
4) W. Güttinger, Phys. Rev. **89** (1953), 1004.

The Meson Theory of Nuclear Forces, II*

— *High Energy Nucleon-nulceon Scattering* —

Saburo FUJII, Junji IWADARE, Shoichiro OTSUKI, Mitsuo TAKETANI†,

Smio TANI†† and Wataro WATARI††

Department of Physics, Kyoto University
† Department of Physics, Rikkyo University
†† Faculty of Engineering, Kyoto University

Using the potentials derived from $Ps(fv)$ meson theory of the second and fourth order, neutron-proton and proton-proton scattering at 40 Mev and 90 Mev in the laboratory system are examined. Calculation is performed by numerical integrations. The characteristic features of high energy nucleon-nucleon scattering are fairly well explained by our pseudoscalar meson potential.

§ 1. Introduction and summary

In the first part of this paper[2]** the deuteron ground state and low energy neutron-proton scattering have been investigated according to the method for treating the problems of nuclear forces proposed by one of us (M. T.) and others[3]. In this paper second and fourth order potentials derived from the pseudoscalar (pv)*** meson theory in the static approximation[4] were adopted near and outside the meson Compton wave length, i.e., in the outside region. In the inside region where the static meson potential becomes meaning-less, phenomenological potential represented by the square well or hard core was adopted. Taking the coupling constant $g^2/4\pi$ between π-meson and nucleon as large as about 0.08, experimental data were fairly well accounted for.

The high energy nucleon-nucleon scattering has been treated by several authors[5]. However, these investigations are almost phenomenological, and their main object is the reproduction of the experimental data on the basis of the assumed potential. So far, the explanation of the high energy nucleon-nucleon scattering data according to the meson theory has not yet been examined sufficiently. Reasons why the sufficient treatment has not been made may be as follows: First, nucleon-nucleon potentials already derived from the meson theory are mainly of the second and fourth order. But in the course of deri-vation of these potentials, many approximations were unavoidably made and their effects

---

* Preliminary report has been published in this journal[1].
** Hereinafter referred to as I.
*** In this paper, (ps) means "with the pseudoscalar coupling" and (pv) means "with the pseudo-vector coupling".

have not been fully estimated.   Moreover, we know little even about the main character of the sixth or eighth order term.   Therefore, the meson potential does not seem to have been finally determined.   Second, when we treat high energy problems, many complicated effects e.g. the velocity dependence of potentials or the relativistic effects, become important. These effects are also scarcely examined.

Recent calculation by Machida and Semba[5] seems to show that the higher order terms do not alter the main characters of the pseudoscalar (ps) potential of the second and fourth order so severely as far as the outside region is concerned.   According to their results, the main terms of the sixth order potential are of the same order of magnitude as the second and fourth order potential where the relative distance between two nucleons is about 0.6 times the meson Compton wave length and are rapidly decreasing in the outside region. In this respect, the recent analysis of the contribution of higher order effects performed by Brueckner* and Watson[7] is also to be noticed.   They have estimated the correction due to the multiple scattering effects to be less than 30% of the results without taking it into account at the distance 0.6 times the meson Compton wave length and this correction is rapidly decreasing outside.

It is true that the value of the coupling constant is not finally determined and there is a room left for the relative weight of the second and fourth order potential to be varied, so that one cannot say that he has a definite answer concerning the problem of the outside potential.   Moreover, the damping effect caused by the terms of higher order in the coupling constant is not thoroughly explored in the pseudoscalar meson theory, and some change may be possible in the conclusion already obtained.   However, one can suppose that he has some pictures about the outside potential of nuclear force which is derived from the pseudoscalar meson theory.   Recent analyses seem to agree in showing similar conclusion about the outside potential of the fourth order both by (ps) theory and by (ps) theory, owing to some damping effects of core terms in (ps) theory.   And, as already mentioned, the (ps) static potential (of the second plus fourth order with $g^2/4\pi = 0.08$) fits well to the low energy data if suitable procedure of cutting off the inside potential is adopted.

Here the results of the investigation on the nucleon-nucleon scattering by several different phenomenological potential[9] (performed by one of the authors (W. W.) and his collaborator) are noticable.   According to them one can safely say that in the case of the collision energy lower than 100 Mev in the laboratory system, S-wave (and sometimes P-wave in the case of very strongly attractive potential) is affected by the change of the inside region of the potential, whereas P-wave and waves with higher angular momentum are mainly scattered by the outside region of the potential.   Then, for the scattering problems up to 100 Mev, the behaviors of P- and higher waves are substantially determined by the outside region of the potential which is given from the meson theory, and these waves play an important role, especially, to determine the angular distribution.   Therefore,

---

* We are indebted to Professor Brueckner for the discussion on this point while he stayed at Kyoto in Sept. 1953.

we think that it is very important at the present stage of meson theory to compare the nucleon-nucleon scattering calculated using the meson theoretical potential with the experimental results up to energy of 100 Mev. For higher energy than 100 Mev, even D-wave is affected by the inside phenomenological potential about which we can say nothing definitely. Therefore, the problems at this energy region are to be treated from different points of view, for example, taking into account the isobar effect or energy dependence of the inside potential.

In this paper, using the potential given in I, the high energy (actually 40 Mev and 90 Mev in the laboratory system) neutron-proton and proton-proton scattering are treated. This potential reproduces the characteristic features of these scattering fairly well. In section 2 the values of phase shifts are given in the Tables, and the angular distributions using these phase shifts are shown. Some corrections to the results of I are made for the parameters of the inside potential of the triplet even state. Discussions about the characteristics of the pseudoscalar meson potential are made in section 3. In section 4 speculation to the future problems is given.

§ 2. Results

In this paper, x is the distance between two nucleons in the unit of meson Compton wave length.

The values of the eigen phase shifts are shown in Tables I and II. For the notations and their meanings, see reference 9. Calculations are performed by the direct numerical integration with the interval Δx and Δy, where the transformation $y = \log x$ is used in the region $x \lesssim 1$, (see I). Low energy parameters given by this potential are as follows :**

Triplet n-p scattering length : $^3a = 5.43 \times 10^{-13}$cm.

Triplet n-p effective range : $^3r = 1.47 \times 10^{-13}$cm.

Binding energy of the deuteron : $|\epsilon| = 2.23$ Mev.

Quadrupole moment of the deuteron : $Q = 2.3 \times 10^{-27}$cm^2.

D-state probability of the deuteron : $p_D = 0.076$.

Singlet n-p scattering length : $^1a = -23.7 \times 10^{-13}$cm.

Singlet n-p effective range : $^1r = 2.26 \times 10^{-13}$cm.*

Some of the values in the Table 3 of I are not correct and our phenomenological inside potential in the triplet even state is different from that adopted in I. The central potential in the triplet even state by the pseudoscalar (ps) meson theory of the second plus fourth order is strongly repulsive, and it exceeds the tensor potential at $x \sim 0.68$ in its absolute value, so that for the deuteron to be bound we have to cut off this central potential at $x \lesssim 0.5$. One example of such phenomenological inside potentials

---

* This value is taken from reference I by interpolation.

** For the experimental values, see I.

is given in Table I. But this may not be the best one and another choice of parameters might fit better to the experiments. We notice that the main part of the sixth order pseudoscalar (ps) meson potential calculated by Machida and Semba[6] is similar to our phenomenological inside potential in the features to decrease the strong repulsive central potential of the second and fourth order. After we finished the above calculation, we were informed that Brueckner and Watson[7] treated the nonadiabatic effect in somewhat different way from the usual perturbation method and obtained another kind of potential in the triplet even state. This potential is similar to that of I but the attractive central force. This fact seems to be favourable to our results in the following sense. In the above calculation we needed to add the correction term in the inside region in order to cancel the effect of the repulsive central force. Concerning this central force Sawada pointed out that the repulsive force comes from inadequacy of the usual perturbation method and the attractive force can be derived not only by Brueckner and Watson's method but also in somewhat different way.[10]

Table I. Triplet eigen phase shift.

State ^3L_J	90 Mev	40 Mev	Computational method	Inside potential
$J=1$ { a-wave (→$^3S_1$ in the absence of tensor force.)	62.0°		numerical integration(N. I.) with $\Delta x=0.1$, $\Delta y=0.05$.	at $x \leq 0.606$, for central potential attractive square well cut off of the depth of 124.6 Mev, for tensor potential zero cut off.
γ-wave (→$^3D_1$)	-2.1°			
amount of admixture η_1	0.136			
$^3D_2$	15.7°		N.I. $\Delta x=0.2$, $\Delta y=0.1$.	
$^3D_3$	-1.0°*		N.I. $\Delta x=0.2$, $\Delta y=0.1$.	
$^3P_0$	21.6°	17.4°	N.I. $\Delta x=0.2$, $\Delta y=0.1$.	
$^3P_1$	-10.7°	-7.0°	N.I. $\Delta x=0.2$, $\Delta y=0.1$.	
$J=2$ { a-wave (→$^3P_2$)	3.7°	1.7°*	N.I. $\Delta x=0.2$, $\Delta y=0.1$.	hard core cut off at $x \leq 0.333$.
γ-wave (→$^3F_2$)	-2.0°			
amount of admixture η_2	0.850			
$^3F_3$	-1.5°		Born approximation.	
$^3F_4$	0.2°		Born approximation.	

* This phase shift is calculated using equivalent potential.[4a]

Table II. Singlet phase shift.

State ^1L	90 Mev	40 Mev	Computational method	Inside potential
^1S	47.3°	64.8°	numerical integration (N.I.) with $\Delta x=0.1$, $\Delta y=0.05$.	hard core at
^1D	2.7°	1.0°	N.I. $\Delta x=0.2$, $\Delta y=0.1$.	$x \leq 0.333$.
^1P	-14.5°	-12.6°	N.I. $\Delta x=0.2$, $\Delta y=0.1$.	hard core at
^1F	-2.2°		Born approximation.	$x \leq 0.333$.

The angular distributions of n-p and p-p scattering in the centre of mass system are shown in Figs. 1, 2, and 3. The total cross section of n-p scattering at 90 Mev is 10.7

$\times 10^{-26}$cm² which is larger compared with the experimental one. As to this point we shall discuss later. The p-p scattering due to Coulomb force is also included but the Coulomb modification of the nuclear phase shift is made only for singlet S-wave. For other waves this modification is negligibly small compared with the nuclear phase shifts.

Fig. 1. n-p scattering at 90 Mev.

————theoretical differential cross section
·······triplet part of the theoretical differential cross section
×   experimental data[11]
—·   experimental data[12]
These two experimental data are normalized to the experimental total cross section of $(7.9 \pm 0.7) \times 10^{-26}$ cm².
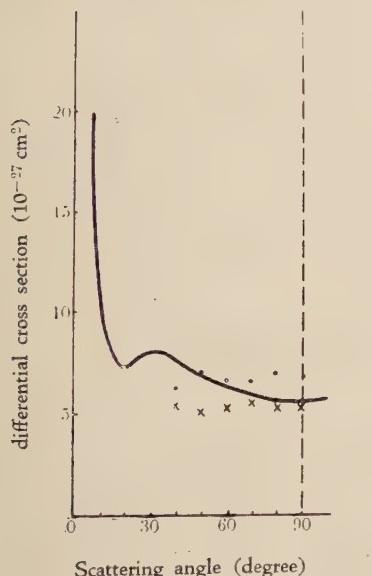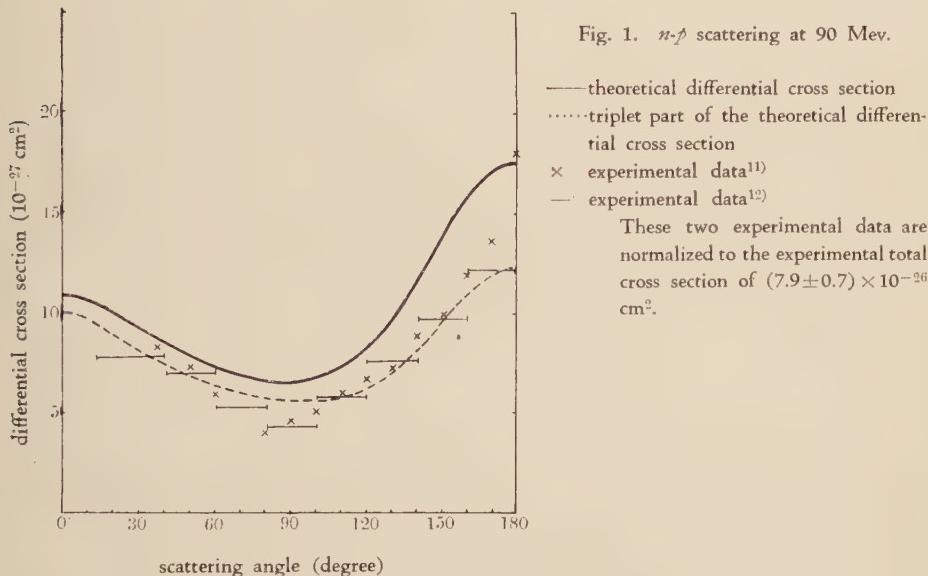
Fig. 2. p-p scattering at 90 Mev.
.   experimental data at 75 Mev[13]
× experimental data at 105 Mev[13]
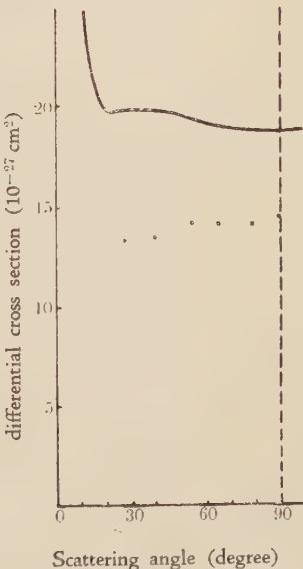
Fig. 3. p-p scattering at 40 Mev.
.   experimental data at 31.8 Mev[14]

§ 3.   Discussions

From the results given in the last section, one can say that the potential given by the pseudoscalar meson theory can give the characteristic features of neutron-proton and proton-proton scattering up to energy of 100 Mev.   The reasons for this statement are as follows :

i)   As has already been shown in I and by Brueckner and Watson[7], it is quite possible for pseudoscalar meson potential to reproduce the low energy scattering data.   This is very satisfactory, because with the appropriate inside potential, S-wave phase shifts at high energies are possibly not as much different from those due to many phenomenological potentials that can give the explanation of high energy nucleon-nucleon scattering data.

As to our results, the total cross section of n-p scattering at 90 Mev is 10.7×10^{-26} cm^2, which is larger than the experimental cross section of $(7.9 \pm 0.7) \times 10^{-26}$ cm^2.[11]   Also the p-p differential cross section at 40 Mev is too large.   These two facts are due respectively to the large phase shift of $^3S_1$-and ^1S-wave.   Here, it is interesting to note that these phase shifts depend on the inside potential sensitively and have some phenomenological character[3].   So that they may be more or less reduced by taking another way of cutting off.

ii)   For the neutron-proton scattering, the meson potential in the triplet odd state is, near and outside its range, not large.   Moreover, tensor potential is stronger than central one in this region.   Therefore, the effects of $^3P_0$-, $^3P_1$-, and $^3P_2$-phase shifts cancel out as a whole, consequently the angular distribution by the triplet potential is almost symmetric about 90°.   Quantitatively, if we expand the triplet differential cross section into the power series of cosine of scattering angle θ as follows[9] :

$$^3\sigma(\theta) = 4/k^2 \cdot \sum c_n \cos^n \theta,$$

where k is the wave number, then the coefficient c_1 is due to the interference of P-waves and other waves.   The ratio $c_1/c_0 = 0.116$ at 90 Mev is very small compared with $c_2/c_0 = 0.315$. In Fig. 4 the potentials effective for $^3P_0$-, $^3P_1$- and $^3P_2$-wave and the wave functions are plotted.   If we denote the central potential in the triplet odd state as V_c and tensor potential as V_t (in the unit of (meson mass)$^2 \times$ (light velocity)^2 / (nucleon mass)), the potential effective for $^3P_0$-state is

$$V_c - 4V_t,$$

and for $^3P_1$-state is

$$V_c + 2V_t.$$

For $^3P_2$-state, the equivalent potential is (without centrifugal force)

$$V_c - (4/5)V_t + 5/x^2 - \sqrt{(25/x^4) - (4/x^2)V_t + (44/5)V_t^2}$$

$$\sim V_c - (2/5)V_t - (22/25)x^2 V_t^2 \qquad \text{for large } x,$$

and the corresponding phase shift is 1.6°.   (We did not use this phase shift to determine the angular distribution, but used the exact phase shifts.   See Table I.)   From Fig. 4, it can be seen that for P-waves the outside potential does almost determine the sign and magnitude of the phase shift,

The large differential cross section in the backward direction is due to the repulsive potential in the singlet odd state, which is also predicted by pseudoscalar (ps) meson theory.[7] This tendency is in agreement with the recent experiment at 135 Mev reported by Snowden at Birmingham conference.*

plot of the potential effective to $^3P_0$-state – – – –
$^3P_1$-state –·–··–·–
$^3P_2$-state ··········

plot of the wave function of $^3P_0$-state – – – – –
$^3P_1$-state –·–··–·–
$^3P_2$-state ··········

plot of the free wave function ———

(Orbinate is in arditrary unit.)
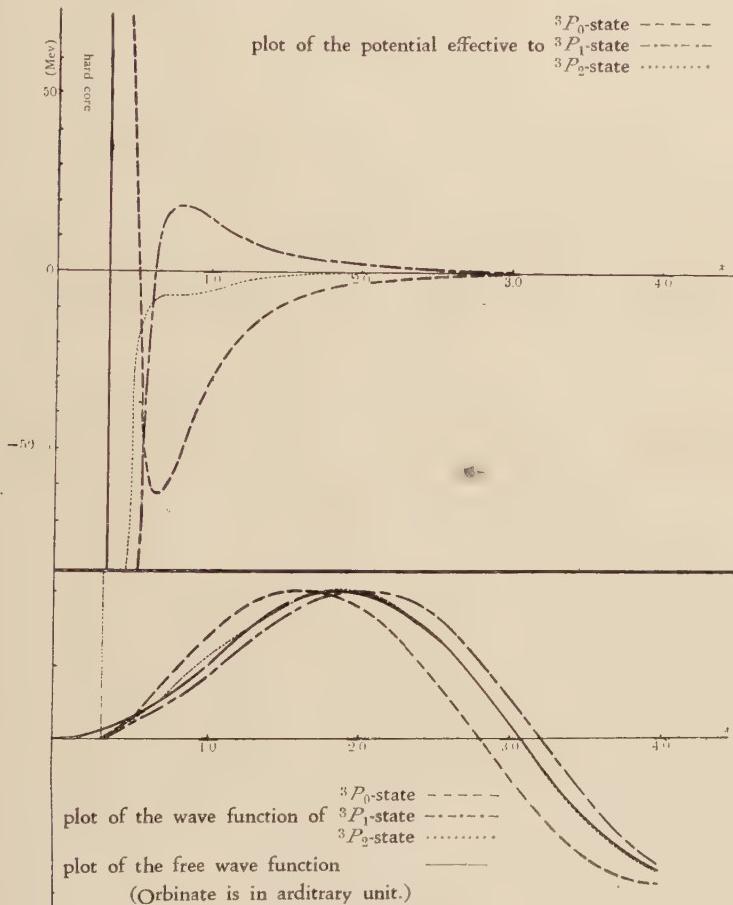
Fig. 4.   plots of the potentials and wave functions in
the triplet P-states.

iii)   For the proton-proton scattering, the potential in the singlet even state is, near and outside its range, very small. Therefore, the phase shift of ^1D-wave is also small and consequently the destructive interference does not break the isotropy of angular distribution so severely.

The polarization $P(\theta, \varphi)$ appearing after p-p scattering by this potential is calculated at 90 Mev, where φ is the azimuthal angle, the z-axis being parallel to the incident

* held at Birmingham in July, 1953,

direction of the beam and x-axis lying in the plane determined by the incident and scattering direction. The polarization of the beam scattered with the angle $\varphi = 0°$ and $\theta \sim 45°$, where $P(\theta, \varphi)$ has its maximum value, is only about 6% and so it is now probably impossible to detect it by the double scattering experiment at this energy.

§ 4. Speculation to the future problems

From the discussions above, we can conclude that pseudoscalar meson potential has a satisfactory characters to explain the experimental data in its outside region, in other words, for P- and higher waves. Of course, many but small corrections will necessarily be added to this outside potential in future from meson theoretical point of view. But at the present stage of meson theoretical approach to nuclear force, we consider it also necessary to find the aspects of the corrections required by experimental data dealing with S-wave phase shifts phenomenologically.

When the appropriate inside potentials are adopted to give the good values of low energy parameters (including deuteron parameters) together with the outside potential derived from meson theory in th region $x \gtrsim 0.6$, can we get high energy ${}^3S_1$- and ${}^1S$-wave phase shifts that reduce the too large cross sections? This is the first problem. It is to be noted that usual phenomenological potentials predicted also larger n-p cross sections.

The potential in the triplet odd state is very delicate : Owing to the cancellation of many terms derived from meson theory, it is very small near the range and its tensor part changes its sign at $x \sim 0.6$. Therefore, even small correction can change the character of the potential features near its range.* For example, the potential of ours and of Brueckner and Watson's[7] agree well in gross features, but the slight difference between them may yield appreciably different values of ${}^3P_0$-, ${}^3P_1$-, and ${}^3P_2$-phase shifts, which are very important to determine both the n-p and p-p angular distribution and p-p polarization. This is not the case for other states. So we must reexamine triplet odd state and, if possible, find the desired potential shape from the experiments. This is the second problem.

---

* This point is first suggested by Mr. S. Machida. We are very much indebted to him for his discussion about the meson theoretical potentials.

References

1) S. Fujii, J. Iwadare, S. Otsuki, M. Taketani, S. Tani and W. Watari, Prog. Theor. Phys. **10** (1953), 478 (L).

2) M. Taketani, S. Machida and S. Onuma, Prog. Theor. Phys. **7** (1952), 45.

3) M. Taktaeni, S. Nakamura and M. Sasaki, Prog. Theor. Phys. **6** (1951), 581.

4) See also K. Nishijima, Prog. Theor. Phys. **6** (1951), 815, 911.

5) For example :
   a) R. C. Christian and E. W. Hart, Phys. Rev. **77** (1950), 441.
   b) R. C. Christian and H. P. Noyes, Phys. Rev. **79** (1950), 85.
   c) K. M. Case and A Pais, Phys. Rev. **80** (1950), 203.
   d) R. Jastrow, Phys. Rev. **81** (1951), 165.

6) S. Machida and K. Semba, Soryusiron-Kenkyu, (Mimeographed Circular in Japanese) **5** (1953), No. 11, 1310.

7) K. A. Brueckner and K. M. Watson, Phys. Rev. **92** (1953), 1023.

8) M. Matsumoto and W. Watari, Prog. Theor. Phys. **11** (1953), 63.

9) F. Rohrlich and J. Eisenstein, Phys. Rev. **75** (1949), 705.

10) K. Sawada, Lecture in the International Conference on Theoretical Physics, Kyoto, Japan, Sept. 1953.

11) Hadley, Kelly, Leith, Segrè, Wiegand and York, Phys. Rev. **75** (1949), 351.

12) Brueckner, Hartsough, Hayward and Powell, Phys. Rev. **75** (1949), 555.

13) R. W. Birge, L. E. Kruse and N. F. Ramsey, Phys. Rev. **93** (1951), 274.

14) Cork, Johnston and Richman, Phys. Rev. **79** (1950), 71.

20

Interaction between Electrons
in One-dimensional Free-electron Model
with Application to Absorption Spectra of Cyanine Dyes

Gentaro ARAKI and Huzihiro ARAKI

Faculty of Engineering, Kyoto University

The interaction between electrons in a deep square-well potential field is discussed in case of one dimension. The potential of one-dimensional two-electron forces is assumed to be represented by a transversal average of the three-dimensional Coulomb potential. The calculated excitation energy is compared with observed data of absorption maxima for cyanine dyes.

Introduction

The one-dimensional free-electron model was first adopted by Baylis[1], Kuhn[2] and Simpson[3] to account for the vinylene shift of absorption maxima of long conjugated molecules. The interaction between electrons have never been taken into account in the theory of the free-electron model. One of the present authors[4] showed, on the basis of the Tomonaga model, that the interaction is very important in the case of long carotenoids. Therefore it must be necessary to examine the interaction in the case of the free-electron model too. This examination is the aim of the present paper. The result of calculation for excitation energy is compared with observed absorption maxima of cyanine dyes with a few vinylene groups. In such a case it is expected from the result of the above-mentioned theory[4] that the interaction is weak. The perturbation method is therefore applied to the calculation in the present paper.

§ 1. First excitation energy

We consider a N-electron system in one dimension, N being even. The electrons are assumed to be enclosed in a very deep potential field of square-well shape. If we neglect the interaction between them the eigenfunction of the system is represented by a determinant of one-electron wave functions. The one-electron states can be characterized by the parity because the potential has a centre of symmetry. The one-electron orbitals and levels are given by[5]

$$\varphi_n(x) = \sqrt{2/L} \sin(\pi nx/L),$$
$$\varepsilon(n) = \{\hbar^2/(2m)\} (\pi/L)^2 n^2 \qquad n = 1, 2, 3, \cdots \qquad (1 \cdot 1)$$

where L is the width of the square well,

In the ground state of the system the orbitals from the first up to the n_0-th are all occupied by electrons where n_0 is equal to $N/2$. In order to satisfy the Pauli principle the ground state should be singlet. In the first singlet excited state an electron in the n_0-th orbital is excited to the (n_0+1)-th orbital. The energy eigenvalues of these states can be calculated by the standard method[1]. The energy difference between these two states is the first excitation energy of the system. If we denote it by $\varDelta E$ it is given by

$$\varDelta E = \varepsilon(n_0+1) - \varepsilon(n_0) + W_c + W_{ex}. \tag{1·2}$$

Here W_c and W_{ex} denote the Coulomb and exchange energies respectively. They are given by

$$W_c = \sum_{n=1}^{n_0-1} 2\{C(n_0+1, n) - C(n_0, n)\} + C(n_0+1, n_0) - C(n_0, n_0),$$

$$W_{ex} = \sum_{n=1}^{n_0-1} \{D(n_0, n) - D(n_0+1, n)\} + D(n_0+1, n_0), \tag{1·3}$$

where C and D are respectively the direct and exchange integrals of the two-electron interaction potential. The calculation of them will be explained in the next section. The one-dimensional two-electron interaction is assumed to be represented by a two-dimensional average of the three-dimensional Coulomb interaction over a transversal region with an area of A. On the basis of this assumption we have

$$\varDelta E = \{\hbar^2/(2m)\}(\pi/L)^2(2n_0+1)(1+\delta), \tag{1·4}$$

where δ is given by (a is the Bohr radius)

$$\delta = (L/\pi)^3(aA)^{-1} 2\{1 - B(n_0)\}(2n_0+1)^{-1},$$

$$B(n) = \frac{5}{8n^2} + \frac{1}{4n(n+1)} - \frac{1}{8(n+1)^2} - \frac{1}{2(2n+1)^2} + \frac{4}{\pi^2}\{1 - G(n)\},$$

$$G(n) = \frac{1}{4(n+1)^2} + \frac{1}{4(n+1)^3} + \frac{1}{(2n+1)^2} - \frac{1}{2(2n+1)^4}$$
$$+ \frac{1}{4}\left\{\frac{1}{n^2} - \frac{1}{(n+1)^2}\right\}\sum_{k=0}^{n-1}\frac{1}{(2k+1)^2} + \frac{1}{4}\left\{\frac{1}{n^3} - \frac{1}{(n+1)^3}\right\}\sum_{k=0}^{n-1}\frac{1}{2k+1}. \tag{1·5}$$

The effect of interaction is represented by δ. Its main part comes from the exchange integrals.

§ 2. Calculation of interaction energy

As is well known, the Culomb interaction potential between two electrons is approximately expanded in the Fourier series in a parallelepiped with the edge lengths of $2L_1$, $2L_2$, and $2L_3$ as follows :

$$\frac{e^2}{r} = J_0 + \frac{\pi e^2}{2L_1 L_2 L_3} \sum_{k \neq 0} \frac{1}{k^2} \exp(i \boldsymbol{k} \boldsymbol{x}), \qquad (2 \cdot 1)$$

where r and $\boldsymbol{x}$ are respectively the distance between and the difference of position vectors of two electrons, and the vector denoted by $\boldsymbol{k}$ is given by

$$k_j = (\pi/L_j) n_j; \quad (j=1, 2, 3; \; n_j=0, \pm1, \pm2, \pm3, \cdots; \; \boldsymbol{k} \neq 0). \quad (2 \cdot 2)$$

If coordinates of two electrons are denoted x_1, y_1, z_1 and x_2, y_2, z_2, eq. $(2 \cdot 1)$ is correct in a domain given by $-L_1 < x_1 - x_2 < L_1$, $-L_2 < y_1 - y_2 < L_2$, and $-L_3 < z_1 - z_2 < L_3$, that is, $0 < x_1, x_2 < L_1$, $0 < y_1, y_2 < L_2$ and $0 < z_1, z_2 < L_3$.

An average of the right side of $(2 \cdot 1)$ over a domain restricted by $-L_2 < y_1 - y_2 < L_2$ and $-L_3 < z_1 - z_2 < L_3$ is given by

$$J(x_1 - x_2) = J_0 + e^2 L(\pi A)^{-1} \sum_{n=1}^{\infty} n^{-2} \cdot \cos(\pi/L) n (x_1 - x_2), \qquad (2 \cdot 3)$$

where L and A stand for L_1 and $L_2 L_3$ respectively. We assume that $J(x_1 - x_2)$ is the two-electron interaction potential in the one-dimensional free-electron model. The direct and exchange integrals in $(1 \cdot 3)$ are then given by

$$C(n, m) = \int_0^L \int_0^L [\varphi_n(x_1) \varphi_m(x_2)]^2 J(x_1 - x_2) dx_1 dx_2,$$

$$D(n, m) = \int_0^L \int_0^L \varphi_n(x_1) \varphi_m(x_1) \varphi_n(x_2) \varphi_m(x_2) J(x_1 - x_2) dx_1 dx_2, \qquad (2 \cdot 4)$$

where $\varphi_n(x)$ is given by $(1 \cdot 1)$. These integrals can be evaluated in the elementary way. The result contains the series of the types given by the left sides of the following equations:

$$\sum_{k=0}^{\infty} \frac{1}{(2k+1)^4} = \frac{1}{6} \left(\frac{\pi}{2}\right)^4,$$

$$\sum_{k=0}^{\infty} \frac{1}{(2k+1)^2} \cdot \frac{1}{(2n)^2 - (2k+1)^2} = \frac{1}{8} \left(\frac{\pi}{2n}\right)^2; \quad n=1, 2, 3, \cdots,$$

$$\sum_{k=0}^{\infty} \frac{1}{(2n)^2 - (2k+1)^2} \cdot \frac{1}{(2m)^2 - (2k+1)^2} = \frac{\delta_{nm}}{16} \left(\frac{\pi}{2n}\right)^2; \quad n, m=1, 2, 3, \cdots, \quad (2 \cdot 5)$$

$$\sum_{k=1}^{\infty} \frac{1}{(2n+1)^2 - (2k)^2} \cdot \frac{1}{(2m+1)^2 - (2k)^2} = \frac{\delta_{nm}}{16} \left(\frac{\pi}{2n+1}\right)^2 - \frac{1}{2(2n+1)^2(2m+1)^2};$$

$$n, m=0, 1, 2, 3, \cdots.$$

These series can be evaluated by making use of the Fourier and Parseval theorems. The Fourier series of the functions $|x|$, $\sin 2n |x|$, and $\sin(2n+1) |x|$ are given by

$$|x| = \frac{\pi}{2} - \frac{4}{\pi} \sum_{k=0}^{\infty} \frac{\cos(2k+1)x}{(2k+1)^2},$$

$$\sin 2n\,|\,x\,| = \frac{4}{\pi}(2n)\sum_{k=0}^{\infty}\frac{\cos(2k+1)x}{(2n)^2-(2k+1)^2}\ ;\ n=1,2,3,\cdots,\qquad (2\cdot6)$$

$$\sin(2n+1)\,|\,x\,| = \frac{4}{\pi}(2n+1)\left\{\frac{1}{2(2n+1)^2}+\sum_{k=1}^{\infty}\frac{\cos 2kx}{(2n+1)^2-(2k)^2}\right\}\ ;\ n=0,1,2,\cdots.$$

These series converge to the functions on the left sides in a domain restricted by $-\pi\leq x\leq\pi$. If we consider the norms of these functions and their Hermitian inner products on this domain we obtain eqs. $(2\cdot5)$ by dint of the Parseval theorem. If we make use of eqs. $(2\cdot5)$ we have the direct and exchange integrals of $J(x_1-x_2)$ as follows :

$$C(n,m)=J_0+\frac{\pi e^2 L}{24A}+\frac{e^2 L}{8\pi A}\left(\frac{1+\delta_{nm}}{n^2}+\frac{1}{m^2}\right);\ n,m=1,2,3,\cdots,$$

$$D(n,m)=\frac{e^2 L}{2\pi A}\left\{\frac{1}{(n-m)^2}+\frac{1}{(n+m)^2}\right\};\ n+m=\text{even}\ ;\ n\neq m\ ;\ n,m=1,2,3,\cdots,$$

$$D(n,m)=\frac{e^2 L}{2\pi A}\left\{\frac{1}{(n-m)^2}+\frac{1}{(n+m)^2}\right\}-\frac{2e^2 L}{\pi^3 A}\left\{\frac{1}{(n-m)^2}-\frac{1}{(n+m)^2}\right\}^2;$$

$$n+m=\text{odd}\ ;\ n\neq m\ ;\ n,m=1,2,3,\cdots.\qquad (2\cdot7)$$

The result given by $(1\cdot4)$ and $(1\cdot5)$ is obtained by substituting these expressions for C and D in $(1\cdot3)$ and thus obtained results for W_c and W_{ex} in $(1\cdot2)$.

Fig. 1

§ 3.   Application to absorption spectra of cyanine dyes

We shall consider a series of cyanine dye cations shown in Fig. 1. In these cations two N-ethylbenzothiazole rings are combined by vinylene groups. The number of vinylene groups will be denoted by v. (An example of $v=3$ is shown in Fig. 1.) Fig. 1 shows only one of main bond structures[4] contained in the state of the cyanine dye cation. The structure, which is obtained by a reflexion regarding a plane through a centre of vinylene bridge, is also contained. If we consider a chain of conjugated double bonds including the vinylene groups and two nitrogen atoms the field for π-electrons contained in the chain is thus symmetrical. Only one double bond of benzene rings is counted in the chain. The chain involves $v+3$ conjugated double bonds, two nitrogen atoms and $2v+5$ carbon atoms. The number of π-electrons contained in these double bonds is equal to $2(v+3)$. If l is the mean length occupied by an atom the length of the chain is equal to $(2v+7)l$.

Now we apply the free-electron model of N electrons considered in the preceding two sections to the system of these π-electrons. We have then $N=2(v+3)$, $n_0=v+3$, and $L=(N+1)l$. We assume, for the sake of simplicity, that l is equal to $1.4\,\text{Å}$ ($=2.646$

atomic units),[4] although the observed length of $C = N$ bond in pyridine is equal to 1.37 Å. The above-mentioned value of l is the bond length in benzene. The absorption maxima calculated on these assumptions are compared with experiment[7] in Table 1. In order to agree with experiment we have to assume the value of $.l$ to be equal to 1000 atomic units. This value is unreasonably large, but it has no real meaning. It only means the smallness of the interaction between electrons. This unreasonable value shows that our model is not quite adequate. It may be partly due to neglect of the variation of wave functions in the transversal dimensions.

Table 1. Absorption maxima of cyanine dyes

n_0 = Fermi max. = number of conj. double bonds
$N = 2n_0$ = number of π-electrons
$N - 1$ = number of carbon atoms in the chain
v = number of vinylene groups
λ = absorption maxima in $m\mu$ (solvent: MeOH)

v	n_0	$N = 2n_0$	δ	λ_{calc}	λ_{obs}
0	3	6	0.0321	438	423
1	4	8	0.0550	551	558
2	5	10	0.0835	656	650
3	6	12	0.1177	752	758

If we compare the absorption maximum of a symmetric cyanine dye with that of a carotenoid of the same number of conjugated double bonds we see that the wave length of the latter is far shorter than that of the former. The real reason for this difference has never fully been known. There is at least one difference between cations of cyanine dyes discussed in this section and carotenoid molecules. The formers have a centre of symmetry but the latters don't. On the other hand, the result of the theoretical calculation[4] may be considered as showing that in carotenoids the periodic property of molecular structures is the predominant cause of determining the property of the wave function, while in case of symmetric cyanine dye cations the end barriers of one-electron fields determine the property of the wave function.

References

1) S. Baylis, J. Chem. Phys. 16 (1948), 287.
2) H. Kuhn, J. Chem. Phys. 16 (1948), 840.
3) W. T. Simpson, J. Chem. Phys. 16 (1948), 1124.
4) G. Araki and T. Murai, Prog. Theor. Phys. 8 (1952), 639.
5) L. I. Schiff, *Quantum Mechanics* p. 34 (1949).
6) E. U. Condon and G. H. Shortley, *Theory of Atomic Spectra* p. 169 (1935).
7) L. G. S. Brooker, P. H. Sprague, C. P. Smyth and G. L. Lewis, J. Am. Chem. Soc. 62 (1940), 1116.
   L. G. S. Brooker, G. H. Keyes and W. W. Williams, J. Am. Chem. Soc. 64 (1942), 199.

On the Bose-Einstein Liquid Model for Liquid Helium, I
——— He^3-He^4 Mixtures* ———

Ziro MIKURA

Research Institute for Iron, Steel and Other Metals, Tohoku University

(Received October 15, 1953)

A theory of He^3-He^4 mixtures is developed, in which He^3 is represented by a nondegenerate Fermi gas in a potential well, and He^4 by an assembly of Bose-particles in another potential well, each Bose-particle being considered to have a much larger mass than a He^4-atom and to have an energy gap between the ground state and the lowest excited state. Agreement between theory and experiment concerning the transition temperature and the vapour pressure in solution is satisfactory, if the energy of excitation is assumed to be proportional to the number density of He^4 " particles " in solution. The increase of the velocity of second sound in solution can be accounted for based on this theory under the condition that He^3 particles partake in the motion of the normal part particles, in contrast to Koide and Usui's conclusion. The osmotic pressure, the thermomechanical effect, and the specific heat in solution are also discussed. It is also remarked that the analogue, in solution, of Mendelssohn and Chandrasekhar's experiment would offer the most definite answer to the question which of the two theories is more tenable, de Boer and Gorter's or the present author's. Second sound for dilute solutions near the absolute zero of temperature is also discussed, and it is concluded that the square of the velocity should be there proportional to T and the proportionality coefficient is roughly equal to $R/3$. Theory and experiment are in general agreement also in this case.

§ 1. Introduction

In order to calculate thermodynamic functions of the solution of He^3 in liquid He^4, de Boer and Gorter[1] used the expression for the ideal mixture of classical liquids under Taconis' assumption that He^3 dissolves only in the normal He^4. On the other hand, Heer and Daunt[2] proposed a theory based on quite a different point of view. The latter authors employed a mixture of two model liquids : they regarded the He^3-He^4 mixture as a mixture of an ideal Fermi-Dirac gas in a smoothed potential well $-\chi_3^0$ and of a degenerate ideal Bose-Einstein gas in a smoothed potential well $-\chi_4^0$. The question which theory is more suitable for treating real mixtures has not yet been answered. It is, moreover, to be noted that the model of an ideal B. E. gas in a potential well does not very well represent the actual situations in pure liquid He^4 ; namely, according to it, the transition should be of the third order instead of the second order, and the specific heat and the density of the normal part should obey $T^{3/2}$-law instead of $T^{5.5 \sim 6}$-law. In the present paper, we employ, instead of Heer and Daunt's ideal B. E. gas, the B. E. gas which is composed of " molecules " with larger particle mass than the actual helium atoms and with an

---

energy gap in the lowest portion of the one-particle-energy spectrum. The B. E. gas thus modified, which was first introduced by Bijl, de Boer, and Michels[3] in 1941 and refined later by Toda[4], is known to reproduce fairly well the actual properties of pure liquid He^4. In the mixture, however, the energy gap will be found necessary to depend on the concentration of He^3 in a particular manner while the particle mass remains unaltered. It is shown that Heer and Daunt's theory thus modified* is more successful than de Boer and Gorter's one in accounting for the experiments on the vapour pressure and the related phenomena of He^3-He^4 mixtures. The "modified B. E. liquid" theory enables us to explain also the increase of the velocity of second sound in mixtures under the plausible assumption that He^3 is carried with the velocity of the normal component of He^4. As Koide and Usui[5] have pointed out, de Boer and Gorter's theory, on the contrary, requires a peculiar assumption that He^3 particles are stationary in the second sound wave propagation. In the last portion of the paper, further possible experiments to test which theory represents better the actual entity, are considered.

The derivation of the formulae according to de Boer and Gorter's theory is not fully included in this paper, as it is found in detail in their original.

§ 2. Thermodynamic functions

A. *The "modified B. E. liquid" theory*. The model adopted here for representing the He^4-component in the mixture is a perfect Bose-Einstein gas consisting of such "molecules" that have the particle mass $M_4 = \nu m_4$ and the excitation energy $W = k\Delta$, situated in a smoothed potential well $-\chi_4^0$. Here m_4 is the atomic mass of He^4 and k the Boltzmann constant. ν will be called the mass factor. The free energy of such a model liquid is

$$F_4 = -kTZ + (N_4/\nu) \cdot kT \ln \lambda_4, \qquad (1)$$

where Z is given by

$$Z = \sum_r w_r \ln \left[1 - \lambda_4 \exp\left(-\varepsilon_r / kT\right)\right]^{-1}, \qquad (2)$$

w_r being the weight of the state r, and the energy ε_r being written as

$$\left.\begin{aligned}
\varepsilon_r &= -\chi_4^0 + k\Delta + p^2/2M_4, \quad r = 1, 2, 3, \ldots \\
\varepsilon_0 &= -\chi_4^0.
\end{aligned}\right\} \qquad (3)$$

λ_4 should be determined by the condition for maintaining a constant number of particles. It should be noticed here that the number of particles is taken as N_4/ν rather than N_4, the number of He^4-atoms. This is necessary for the normal fraction and the specific heat of the model liquid to be in accord with the observed values in the actual liquid He^4**. This situation leads one to suppose that, in the liquid state, ν He^4-atoms cluster into a

---

* This theory will be called hereafter the "modified B. E. liquid" theory.
** See the foot note on p. 30.

large " molecule ", rather than to suppose that each atom has a large effective mass like an electron in metal. Introducing (3) in (2) and putting $\lambda_4^* = \lambda_1 \exp(\chi_1'/kT)$, we get

$$Z = -\ln\left[1 - \lambda_4^*\right] - 2\pi\left(2M_4kT/h^2\right)^{3/2}\left(N_3V_3^0 + N_4V_4^0\right)\int_0^\infty \ln\left[1 - \lambda_4^* \exp\left(-\Delta/T - y\right)\right]y^{1/2}dy, \tag{4}$$

where the assumption for an ideal solution is made, that the total volume is given by $V = N_3V_3^0 + N_4V_4^0$, V_3^0 and V_4^0 being the atomic volumes per atom in pure liquid He^3 and He^4 respectively. After integration, we get

$$Z = -\ln\left[1 - \lambda_4^*\right] - \left(2\pi M_4kT/h^2\right)^{3/2}\left(N_3V_3^0 + N_4V_4^0\right)F_{5/2}(\alpha' + \Delta/T), \tag{5}$$

where $\alpha' = -\ln\lambda_4^* > 0$ and

$$F_k(\eta) = e^{-\eta} + e^{-2\eta}/2^k + e^{-3\eta}/3^k + \cdots\cdots . \tag{6}$$

The free energy of a perfect Fermi gas, in a potential well $-\chi_3^0$, with spin $1/2$ is given by

$$F_3 = -kT\left\{4\pi\left(2m_3kT/h^2\right)^{3/2}\left(N_3V_3^0 + N_4V_4^0\right)\int_0^\infty \ln\left[1 + \lambda_3\exp\left(\chi_3^0/kT - y\right)\right]y^{1/2}dy - N_3\ln\lambda_3\right\}, \tag{7}$$

where λ_3 is also the normalizing parameter. The free energy of the total system then becomes

$$F = F_3 + F_4 = -kT\left\{4\pi\left(2m_3kT/h^2\right)^{3/2}\left(N_3V_3^0 + N_4V_4^0\right)\right.$$

$$\left. \times \int_0^\infty \ln\left[1 + \lambda_3 \exp\left(\chi_3^0/kT - y\right)\right]y^{1/2}dy - N_3\ln\lambda_3 - N_4/\nu \cdot \ln\lambda_4 + Z\right\}. \tag{8}$$

Following Heer and Daunt[2], we take as the normalizing condition $(\partial F/\partial\lambda_4)_{T,V} = 0$, which leads to

$$\lambda_4^*/(1 - \lambda_4^*) - N_4/\nu + \left(2\pi M_4kT/h^2\right)^{3/2}\left(N_3V_3^0 + N_4V_4^0\right)F_{3/2}(\alpha' + \Delta/T) = 0. \tag{9}$$

When the temperature is higher than the temperature of condensation, λ_4^* is less than unity and $\lambda_4^*/1 - \lambda_4^*$ may be neglected as compared with N_4/ν. Thus we obtain

$$N_4/\nu = \left(2\pi M_4kT/h^2\right)^{3/2}\left(N_3V_3^0 + N_4V_4^0\right)F_{3/2}(\alpha' + \Delta/T) \qquad T > T_\lambda. \tag{10}$$

On the other hand, for $T \leq T_\lambda$ where $\lambda_4^* \approx 1$, we may put with London[6]

$$\lambda_4^* = 1 - 1/(1 - x) \cdot \nu/N_4, \tag{11}$$

where $x = \rho_{1n}/\rho_4$ is the fraction of the non-condensed particles in the He^4-component. Introducing (11) into (9), we have

$$N_4/\nu \cdot x = \left(2\pi M_4kT/h^2\right)^{3/2}\left(N_3V_3^0 + N_4V_4^0\right)F_{3/2}(\Delta/T). \tag{12}$$

At $T = T_\lambda$, setting $\alpha' = 0$ in eq. (10) gives

$$N_4/\nu = \left(2\pi M_4kT_\lambda/h^2\right)^{3/2}\left(N_3V_3^0 + N_4V_4^0\right)F_{3/2}(\Delta/T_\lambda), \tag{13}$$

which gives the lamda-temperature as a function of He^3-concentration if the dependences

of $\varDelta$ and M_4 on concentration are known. As will be shown below, $\varDelta$ should be considered as dependent on concentration but M_4 (or ν) as independent of it.

At ordinary pressures, F can be used approximately in place of Gibbs' function G. Then the partial potentials are

$$\mu_3 = (\partial G/\partial N_3)_{p,T,N_4} = (\partial F/\partial N_3)_{p,T,N_4},\quad \mu_4 = (\partial G/\partial N_4)_{p,T,N_3} = (\partial F/\partial N_4)_{p,T,N_3}.$$

We now assume that the liquid He^3 model, pure and diluted, is perfectly non-degenerate, namely Boltzmannian.* Then introducing (5) into (8) and differentiating it by N_3 and N_4 respectively, we get the following expressions:

$T \leqq T_\lambda$

$$\mu_3 = -x_3^0 - kT\,\frac{N_3 V_3^0}{N_3 V_3^0 + N_4 V_4^0} + kT\ln\frac{N_3}{N_3 V_3^0 + N_4 V_4^0} - kT\ln 2\Big(\frac{2\pi m_3 kT}{h^2}\Big)^{3/2}$$
$$-kT\Big(\frac{2\pi M_4 kT}{h^2}\Big)^{3/2}\Big\{V_3^0 - (N_3 V_3^0 + N_4 V_4^0)\frac{1}{T}\Big(\frac{\partial \varDelta}{\partial N_3}\Big)_{N_4}\Big\}\exp\Big(-\frac{\varDelta}{T}\Big),\quad (14)$$

$$\mu_3^0 = -\chi_3^0 - kT - kT\ln V_3^0 - kT\ln 2\,(2\pi m_3 kT/h^2)^{3/2},\tag{15}$$

$$\mu_4 = -\chi_4^0/\nu - kT\,N_3 V_4^0/(N_3 V_3^0 + N_4 V_4^0)$$
$$-kT(2\pi M_4 kT/h^2)^{3/2}\{V_4^0 - (N_3 V_3^0 + N_4 V_4^0)1/T\cdot(\partial \varDelta/\partial N_4)_{N_3}\}\exp(-\varDelta/T),\tag{16}$$

$$\mu_4^0 = -\chi_4^0/\nu - kT(2\pi M_4 kT/h^2)^{3/2}V_4^0\exp(-\varDelta_0/T).\tag{17}$$

$T_\lambda^0 \geqq T \geqq T_\lambda$

$$\mu_3 = -\chi_3^0 - kT\,\frac{N_3 V_3^0}{N_3 V_3^0 + N_4 V_4^0} + kT\ln\frac{N_3}{N_3 V_3^0 + N_4 V_4^0} - kT\ln 2\Big(\frac{2\pi m_3 kT}{h^2}\Big)^{3/2}$$
$$-\frac{kT}{\nu}\Big\{\frac{N_4 V_3^0}{N_3 V_3^0 + N_4 V_4^0} - \frac{N_4}{T}\Big(\frac{\partial \varDelta}{\partial N_3}\Big)_{N_4}\Big\},\tag{18}$$

$$\mu_3^0 = -\chi_3^0 - kT - kT\ln V_3^0 - kT\ln 2\,(2\pi m_3 kT/h^2)^{3/2},\tag{19}$$

$$\mu_4 = -\frac{\chi_4^0}{\nu} - kT\,\frac{N_4 V_4^0}{N_3 V_3^0 + N_4 V_4^0} - \frac{kT}{\nu}\Big\{\frac{N_4 V_4^0}{N_3 V_3^0 + N_4 V_4^0} - \frac{N_4}{T}\Big(\frac{\partial \varDelta}{\partial N_4}\Big)_{N_3}\Big\}$$
$$+\frac{kT}{\nu}\ln\frac{N_4/\nu}{N_3 V_3^0 + N_4 V_4^0} + \frac{k\varDelta}{\nu} - \frac{kT}{\nu}\ln\Big(\frac{2\pi M_4 kT}{h^2}\Big)^{3/2},\tag{20}$$

$$\mu_4^0 = -\chi_4^0/\nu - kT/\nu\cdot(2\pi M_4 kM/h^2)^{3/2}V_4^0\exp(-\varDelta_0/T).\tag{21}$$

$T \geqq T_\lambda^0$

μ_3, μ_3^0, and μ_4 have the same forms as (18), (19), and (20).

$$\mu_4^0 = -\chi_4^0/\nu - kT/\nu - kT/\nu\ln\nu V_4^0 + k\varDelta_0/\nu - kT/\nu\cdot\ln(2\pi M_4 kT/h^2)^{3/2}.\tag{22}$$

---

* c. f. Heer and Daunt[2)] and § 10 of this article.

In deriving these equations $F_{3/2}(\omega' + \Delta/T)$ and $F_{5/2}(\omega' + \Delta/T)$ are all replaced by $\exp(-\omega' - \Delta/T)$, the first term in the expansion (6). This replacement is obviously allowed to a good degree of accuracy in the case of $\Delta = \Delta_0 (= 8.609°\mathrm{K})$ and $T = T_\lambda^0 (= 2.186°\mathrm{K})$. This is also true on account of the lowering of the transition temperature for the solution at this temperature, as easily seen from the results of the next section. At temperatures above and below the transition, it is easily found from eqs. (10) and (12) that $F_{3/2}$ is always less than the value at the transition temperature and that in consequence it can be replaced by the exponential function. Since $F_{5/2}$ is always less than $F_{3/2}$ with the same argument, it is concluded that both $F_{3/2}(\omega' + \Delta/T)$ and $F_{5/2}(\omega' + \Delta/T)$ can be safely replaced by $\exp(-\omega' - \Delta/T)$ at all temperatures and concentrations. Introducing (12) into (14), (16), (17), and (21), we get:

$T \leq T_\lambda$

$$\mu_3 = -\chi_3^0 - kT \frac{N_3 I_3^{\prime 0}}{N_3 I_3^{\prime 0} + N_4 I_4^{\prime 0}} + kT \ln \frac{N_3}{N_3 I_3^{\prime 0} + N_4 I_4^{\prime 0}} - kT \ln 2 \left(\frac{2\pi m_3 kT}{h^2} \right)^{3/2}$$
$$- \frac{kT}{\nu} \left\{ C_4' \frac{I_3^{\prime 0}}{I_4^{\prime 0}} - \frac{N_4}{T} \left(\frac{\partial \Delta}{\partial N_3} \right)_{N_4} \right\} x, \tag{14'}$$

$$\mu_4 = -\frac{\chi_4^0}{\nu} - kT \frac{N_3 V_4^0}{N_3 V_3^0 + N_4 V_4^0} - \frac{kT}{\nu} \left\{ C_4' - \frac{N_4}{T} \left(\frac{\partial \Delta}{\partial N_3} \right)_{N_4} \right\} x, \tag{16'}$$

$$\mu_4^0 = -\chi_4^0/\nu - kTx_0/\nu, \tag{17'}$$

where $C_4' = N_4 V_4^{\prime 0}/N_3 V_3^0 + N_4 I_4^{\prime 0}$.

$T_\lambda^0 \geq T \geq T_\lambda$

$$\mu_4^0 = -\chi_4^0/\nu - kTx_0/\nu. \tag{21'}$$

B. *de Boer and Gorter's theory.* de Boer and Gorter assumed Gibbs' energy G of the He^3-He^4 mixture to have the form:

$$G(p,\ T,\ x,\ N_3,\ N_4) = N_4 g_4^0(p,\ T,\ x) + N_3 g_3^0(p,\ T)$$
$$+ kTxN_4 \ln xN_4/(N_3 + xN_4) + kTN_3 \ln N_3/(N_3 + xN_4), \tag{23}$$

where N_1, xN_4 and N_3 are the numbers of atoms of He^4 in superfluid state, He^4 in normal state and He^3 respectively in the solution, and g_4^0 and g_3^0 are the Gibbs free energies per atom of He^4 and He^3 respectively in pure states. Following de Boer and Gorter[1] we put

$$g_4^0(p,\ T,\ x) = -5.5/6.5 \times 2.186 \, m_4 S_\lambda (1 - x^{6.5/5.5}) - m_4 S_\lambda xT, \tag{24}$$

where we employ Tisza's model

$$x_0 = (T/2.186)^{5.5}, \qquad S = S_\lambda x, \qquad S_\lambda = 0.405 \text{ cal/g. deg.} \tag{25}$$

for pure He^4. As He^3, even in the pure state, is considered nondegenerate, the Gibbs free energy per atom is

$$g_3^0 = \text{const.} - kT - kT \ln V_3^0 - kT \ln 2 \, (2\pi m_3 kT/h^2)^{3/2}, \tag{26}$$

§ 3.   The transition temperature and the normal fraction

Heer and Daunt[2] showed that the "ideal B. E. liquid" theory gives the transition temperatures of He^3-He^1 mixtures in moderate agreement with experiment. But, as we shall see below, the revision of their theory by adopting the "modified B. E. liquid" described in the previous section would violate this agreement if the possible dependence of Δ on He^3-concentration were not counted in. A tentative choice of the dependence given by*

$$\Delta = \Delta_0 \, N_4 V_4^0 / (N_3 V_3^0 + N_4 V_4^0) = C_4' \Delta_0, \qquad (27)$$

assuming a constant mass $M_i = M_1^0 = \nu m_1$, will be found to give the transition temperature in fairly good agreement with experiment, at least for concentrations smaller than about 50 percent. Let the concentration of He^3 be denoted by

$$\hat{\xi}_3 = N_3 / (N_3 + N_1),$$

then the transition temperature can be obtained by solving the following simple equation numerically :

$$T_\lambda / T_\lambda^0 = \{(1-\hat{\xi}_3)/1 + (V_3^0/V_4^0 - 1)\hat{\xi}_3 \cdot F_{3/2}(\Delta_0/T_\lambda^0)/F_{3/2}(\Delta/T_\lambda)\}^{2/3}, \qquad (28)$$

which follows immediately from (13). Eq. (28) reduces to Heer and Daunt's equation

$$T_\lambda / T_\lambda^0 = \{(1-\hat{\xi}_3)/1 + (V_3^0/V_4^0 - 1)\hat{\xi}_3\}^{2/3}$$

by setting $\Delta = \Delta_0 = 0$. In order to compare the theory with experiment, the numerical values of parameters, ν and Δ_0, are required. They should be determined properly so that the transition temperature of pure liquid He^4, calculable from (13) with $\Delta = \Delta_0$, may become $2.186°K$ and that all temperature dependence of the normal fraction x_0, the specific heat c_v^0, and the velocity of second sound v_{II}^0 may be in reasonable agreement with experiment**. The values adopted here are

$$\nu = 8.8 \quad \text{and} \quad \Delta_0 = 8.609°K.$$

The normal fraction, the specific heat, and the velocity of second sound, in pure He^4,

Fig. 1. ρ_n/ρ for pure liquid He^4.
- - - - : "Modified Bose-Einstein liquid" model : $\nu=8.8$, $\Delta_0=8.609°K$.
———— : $[x_0]_{obs.}$ derived from observed values of S^0, C_p^0 and v_{II}^0.
$\times$, $+$, $\bigcirc$, $\bullet$: Experimental points by Hollis-Hallett[9], oscillating disc method.

* This relation means that the excitation energy is proportional to the number density of He^4 particles.
** If one replaced the particle number, N_4/ν, on the left-hand side of (12) by N_4, any choice of the parameters would not enable one to make the calculated values of x_0 and c_v^0 fit the experimental data simultaneously.
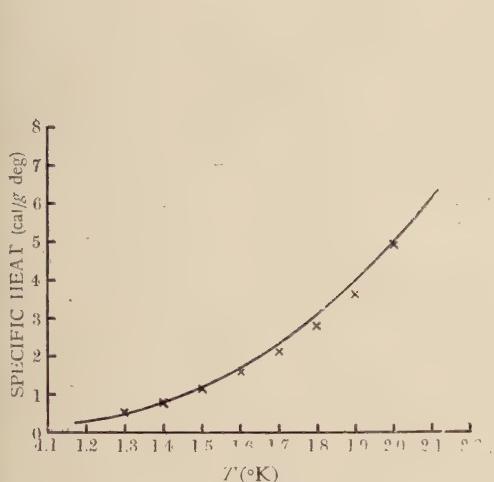
Fig. 2.   Specific heat of pure liquid He⁴.
——: "Modified Bose-Einstein liquid" model:
       $\nu=8.8$, $\Delta_0=8.609°$K.
× : Experimental points, Kramers *et al.*[13]
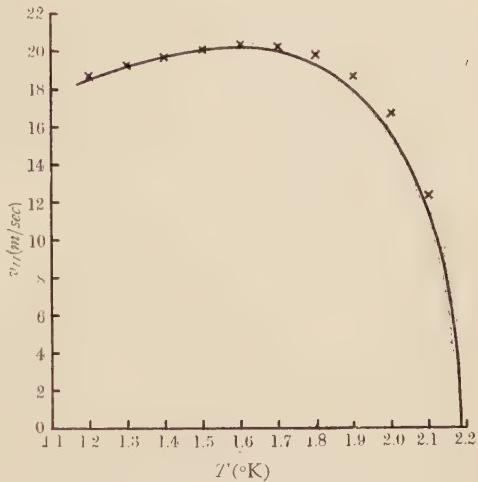
Fig. 3.   Second-sound velocity in pure He⁴.
——: "Modified Bose-Einstein liquid" model:
       $\nu=8.8$, $\Delta_0=8.609°$K.
× : Experimental points, Pellam[14] and Maurer and Herlin.[15]

calculated with these values are shown and compared with experiment in Figs. 1, 2 and 3. The formulae used for computing those values are eq. (12) as well as the following:

$$c_v^0 = \frac{k\Delta_0^2}{\rho_4^0}\left(\frac{2\pi M_4 k}{h^2}\right)^{3/2} T^{-1/2}\left(1+\frac{3}{\Delta_0}T+\frac{15}{4\Delta_0^2}T^2\right)\exp(-\Delta_0/T),$$

$$v_{II}^{02} = (1-x_0)\,TS_0^2/x_0 c_v^0. \tag{29}$$

It is easy to show that eq. (29) for v_{II}^{02} is equivalent to the equation obtained by Gorter, Kasteleijn, and Mellink[7] in the limiting case of low frequency, i.e. to the equation

$$v_{II}^{02} = x_0(1-x_0)\,G_{xx}^0/(1-G_{xx}^0 G_{Tl}^0/G_{xT}^{02}), \tag{30}$$

provided Gorter's condition[8] for H. London's relation to hold in very narrow slits, that is,

$$x_0(\partial S_0/\partial x_0)_{p,T}=S_0 \tag{31}$$

is satisfied. G_{xx}^0, G_{TT}^0 and G_{xT}^0 are the abbreviations for $(\partial^2 G^0/\partial T^2)_{p,T}$, etc.

Employing the empirically known value $l_3^{\cdot 0}/l_4^{\cdot 0}=1.36$, we can compute T_λ/T_λ^0 from (28). The results are shown in Fig. 4, together with the experimental points taken from Heer and Daunt's[2] paper. The agreement is fairly good for solutions containing less than ca. 50 percent He³. The discrepancies between theory and experiment at higher concentrations are not surprising, since the transition temperatures in this case lie below 1°K and the "modified B. E. liquid" model neglecting all influences of phonons should be considered inadequate in this temperature region. In order to exhibit the importance of the dependence of Δ on He³-concentration, the results obtainable when this dependence is not counted in are also shown in Fig. 4 (broken line). The values for T_λ/T_λ^0 calculated by

the de Boer and Gorter method are also included in the same figure.

Now we shall consider the normal fraction x in solution. It may be computed from (12). The results of this computation as well as those obtained from de Boer and Gorter's
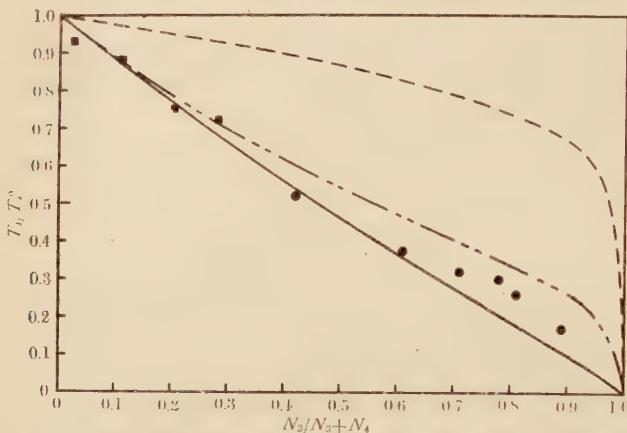
Fig. 4. T_λ/T_λ^0 for solutions of He³ in He⁴.

———— : "Modified Bose-Einstein liquid" theory : $\Delta \propto N_1 V$.

- - - - : "Modified Bose-Einstein liquid" theory : $\Delta =$ const.

—··— : de Boer and Gorter's theory.

■ , ● : Experimental points, taken from Heer and Daunt's paper[2].

theory, are shown in Fig. 5, for solutions containing less than 20 percent He³. Though at temperatures near the transition points both theories give nearly equal values of x, at temperatures well below the transition points de Boer and Gorter's theory gives considerably higher values than the "modified B. E. liquid" theory. Thus, the determination of x in solution by the similar experiment as Hollis-Hallett[9] performed with pure liquid He⁴, will be helpful to decide which is better of the two theories.

The only experiment reported till now, which seems to serve this purpose, is that made by Hammel and Schuch[10]. They measured the critical rates of film flow at various temperatures for a 3.9 percent solution. We can deduce from their data the approximate values of x in the solution relative to those in pure He⁴, if one may assume at least approximately : (1) The critical rates of film flow

Fig. 5. ρ_n/ρ_1 for solutions with He³-concentrations less than 20 percent.

———— : "Modified Bose-Einstein liquid" theory,

—··— : de Boer and Gorter's theory.

are mainly determined by the flow of He⁴ atoms even in solution, at least, provided it is dilute, (ii) the thickness of creeping films does not considerably change with temperature, (iii) the temperature-dependence of the critical rate of flow is of the same form as that of $\rho_s/\rho_1 = 1 - x$, and (iv) the mean particle-velocities at the critical rate of flow are not considerably different in pure He⁴ and in the solution, at least when the latter is dilute.

Among these four assumptions, the first one, (i), may be regarded as an empirical fact ; even in Hammel and Schuch's experiment, in which it was reported that some of He³ did superflow, the contribution of He³-atoms to superflow was small compared with that of He⁴-atoms. Assumptions (ii) and (iii) have been confirmed by experiment for pure He[¹ ¹¹),¹²)]. Their extension to the case of dilute solutions will be probably admissible. From (ii) and (iii) we can then conclude that the mean particle velocity at the critical rate of flow does not change with temperature. Thus the conclusion can be derived from (i), (ii) and (iii) that the ratio of the critical rates of flow at different temperatures approximately represents the ratio of $1-x$ at the corresponding temperatures. If assumption (iv) is also admissible then the above statement has such an universal validity, that the two critical rates of flow concerned can be connected not only with the identical solution but also with the solutions with different concentrations. On the basis of these considerations,

Hammel and Schuch's observed critical rates of flow, R_c, relative to the value for pure He⁴ at $T = 1.96°$K can be compared with theoretical values of $1-x$ relative to the value for pure He⁴ also at $T = 1.96°$K. The results of comparison are shown in Fig. 6. The large discrepancies between theory and experiment, especially at lower temperatures, do not permit us to derive any conclusion from this comparison. At the present stage of experimental study, however, it would not be safe to conclude that one of the four assumptions described is not valid even approximately. It may be doubtful that the He³-concentration in the portion of film where the critical rates of flow were determined, was carefully maintained of the same value as in the bulk liquid in the vessel, for Hammel and Schuch's experiment was not planned for the same purpose as we have quoted it here. It is desirable to carry out in future the more precise experiment on the flow of films with an exactly known He³-concentration.

Fig. 6. Comparison between relative values of theoretical superfluid fraction and experimental critical rate of film flow for the 3.9 percent solution. Both $R_{c,0}$ and x_0 refer to the values for pure He⁴ at 1.96°K.

§ 4.   The vapour pressure

The saturated vapours in contact with the liquid phase at temperatures under consideration are considered as perfect gases. Then the partial vapour pressure of He³ and He⁴ in equilibrium with the solution can be computed from the equations :

$$kT \ln p_3/p_3^0 = \mu_3' - \mu_3^{0\prime} = \mu_3 - \mu_3^0, \tag{32}$$

$$kT \ln p_4/p_4^0 = \mu_4' - \mu_4^{0\prime} = \mu_4 - \mu_4^0, \tag{33}$$

where the primed partial potentials refer to the vapour phase, which must be equal to the partial potentials for liquid phase, and p_3 and p_4 are the partial vapour pressures of He³ and He⁴ of the solution, and p_3^0 and p_4^0 the full vapour pressures of pure liquid He³

and He^4 respectively.

A. *The "modified B. E. liquid" theory.* From eq. (27) we get

$$(\partial J/\partial N_3)_{N_4} = -J \cdot V_3^{-0}/(N_3 I_3^{-0} + N_4 I_4^{-0}), \quad (\partial J/\partial V_4)_{N_3} = N_3/N_4 \cdot J \cdot I_3^{-0}/(N_3 I_3^{-0} + N_4 I_4^{-0}).$$

The partial vapour pressures p_3 and p_4 can be then calculated from (32) and (33) by employing (14′), (15), (16′), (17′), (18), (19), (20), (21′), and (22).

The results are the following :

$T \leqq T_\lambda$

$$p_3 = p_3^0 C_3' \exp\left[C_4' \left\{ 1 - \frac{1}{\nu} \frac{V_3^{-0}}{I_4^{-0}} \left(1 + \frac{\Delta}{T} \right) x \right\} \right], \tag{34}$$

$$p_4 = p_4^0 \exp\left[-\frac{V_4^0}{V_3^0} C_3' - \frac{C_4'}{\nu} \left(x - \frac{x_0}{C_4'} - \frac{N_3}{N_4} \frac{V_3^0 \Delta}{I_4^{-0} T} x \right) \right], \tag{35}$$

where $C_3' = 1 - C_4' = N_3 V_3^0 / N_3 V_3^0 + N_4 V_4^0.$

$T_\lambda^0 \geqq T \geqq T_\lambda$

$$p_3 = p_3^0 C_3' \exp\left[C_4' \left\{ 1 - \frac{1}{\nu} \frac{V_3^{-0}}{I_4^{-0}} \left(1 + \frac{\Delta}{T} \right) \right\} \right], \tag{36}$$

$$p_4 = p_4^0 \left\{ \left(\frac{T_\lambda}{T} \right)^{3/2} \exp\left[\frac{\Delta}{T} - \frac{\Delta}{T_\lambda} \right] \right\}^{1/\nu} \exp\left[-\frac{V_4^0}{V_3^0} C_3' - \frac{C_4'}{\nu} \left(1 - \frac{x_0}{C_4'} - \frac{N_3}{N_4} \frac{I_3^{-0}}{I_4^{-0}} \frac{J}{T} \right) \right]$$

$$= p_4^0 \left\{ C_4' \left(\frac{T_\lambda^0}{T} \right)^{3/2} \exp\left[\frac{\Delta}{T} - \frac{\Delta_0}{T_\lambda^0} \right] \right\}^{1/\nu} \exp\left[-\frac{V_4^0}{V_3^0} C_3' - \frac{C_4'}{\nu} \left(1 - \frac{x^0}{C_4'} - \frac{N_3}{N_4} \frac{V_3^0 \Delta}{V_4^0 T} \right) \right]. \tag{37}$$

$T \geqq T_\lambda^0$

$$p_3 = p_3^0 C_3' \exp\left[C_4' \left\{ 1 - \frac{1}{\nu} \frac{V_3^{-0}}{I_4^{-0}} \left(1 + \frac{\Delta}{T} \right) \right\} \right], \tag{38}$$

$$p_4 = p_4^0 \left\{ C_4' \exp\left[\frac{\Delta}{T} - \frac{\Delta_0}{T} \right] \right\}^{1/\nu} \exp\left[-\frac{V_4^{-0}}{I_3^{-0}} C_3' - \frac{C_4'}{\nu} \left(1 - \frac{1}{C_4'} - \frac{N_3}{N_4} \frac{I_3^{-0}}{I_4^{-0}} \frac{J}{T} \right) \right]. \tag{39}$$

It should be remarked that these formulae are valid at all concentrations and temperatures above 1.3°K, provided that the effect of the non-ideality of the vapour is not serious.

B. *de Boer and Gorter's theory.* Consulting to their paper[1], the partial vapour pressures are given by

$T < T_\lambda$

$$p_3 = p_3^0 \, \hat{\xi}_3 / (x(1 - \hat{\xi}_3) + \hat{\xi}_3), \tag{40}$$

$$p_4 = p_4^0 \left\{ x(1 - \hat{\xi}_3) / x(1 - \hat{\xi}_3) + \hat{\xi}_3 \right\}^x \exp\left[1/kT \cdot \left\{ g_4^0(x) - g_4^0(x_0) \right\} \right]. \tag{41}$$

$T_\lambda^0 \geqq T \geqq T_\lambda$

$$p_3 = p_3^0 \hat{\xi}_3, \quad p_4 = p_4^0 (1 - \hat{\xi}_3) \exp[1/kT \cdot \left\{ g_4^0(1) - g_4^0(x_0) \right\}]. \tag{42}$$

$T \gtrless T_\lambda^0$

$$p_3 = p_3^0 \xi_{33}, \quad p_1 = p_1^0 (1 - \xi_{33}).$$
$$(43)$$

$g_4^0(x)$ has already been given in § 2 (eq. (24)).

C. *Comparison with experiment.* Practical applications of the formule obtained require the precise estimation of the normal fraction x in solution. As is seen in Fig. 1, the curve of x_0 for pure He^4 calculated in the "modified B. E. liquid" theory from eq. (12) does not very well fit the observed values. It is supposed from this fact that the calculation, by the "modified B. E. liquid" theory, of x in solution does not also give so good values as to be applicable for the present purpose. Consequently, we have estimated x in solution in the following way. The values of x_0 for pure He^4 have been first calculated from the empirical data of the entropy S_0,[13] the specific heat c_p^{0} [13], and the velocity of second sound[14], v_{II}^{0} [15] by employing eq. (29), assuming $c_p^0 = c_v^0$. We denote thus obtained values by $[x_0]_{obs.}$ which are in complete accordance with those empirically obtained by Hollis-Hallett[9] from the viscosity measurements by the oscillating disc method. (c.f. Fig. 1) The normal fraction in solution has been then derived by applying the equation

$$x = [x_0]_{obs.} \times (x/x_0)_{theor.} \quad (44)$$

where $(x/x_0)_{theor.}$ is the ratio of x in solution to x_0 in pure He^4, calculated theoretically from (12) and (27).
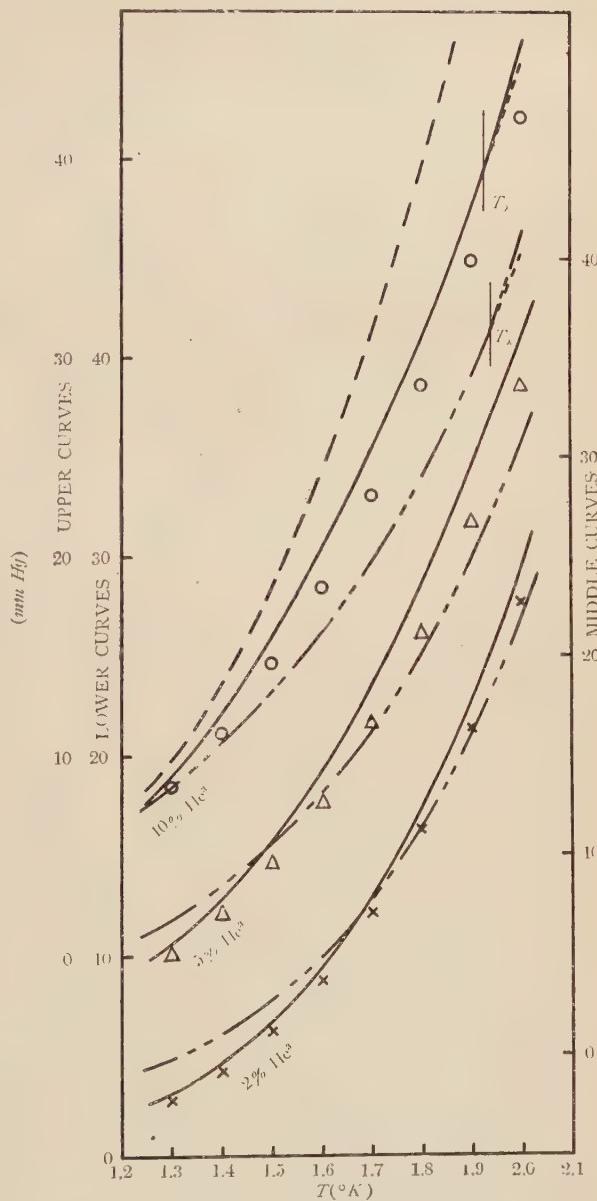
Fig. 7. Relation between vapour pressure and temperature.
———: "Modified Bose-Einstein liquid" theory: $\varDelta \propto N_3/V$.
– – – –: "Modified Bose-Einstein liquid" theory: $\varDelta$=const.
—·—·—: de Boer and Gorter's theory.
Experimental points: Sommers' smoothed data[16]

Theoretical estimations of x in de Boer and Gorter's theory are not corrected, since x_0 in this theory does not deviate so much from observation.

Using these values for x and applying eqs. from (34) to (43), the total vapour pressure $p = p_3 + p_4$ is computed with the 2.0, 5.0 and 10.0 percent solutions. Fig. 7 shows the results of calculation together with the smoothed observed values taken from Table III in Sommers'[16] paper. Both theories are in moderate agreement with observation, though de Boer and Gorter's theory seems to give somewhat too high values for dilute solutions at lower temperatures. It would also be noted that the $p - T$ curve above T_λ in the "modified B. E. liquid" theory as well as in de Boer and Gorter's theory lies above the extrapolated curve from below T_λ, showing a discontinuity in slope at $T = T_\lambda$. (c.f. Fig. 7).

In order to show the importance of the dependence of $\varDelta$ on He^3-concentration, given by (27), is also calculated with the 10 percent solution, assuming $(\partial\varDelta/\partial N_3)_{N_4} = (\partial\varDelta/\partial N_4)_{N_3} = 0$. The results are shown in Fig. 7 by a dashed curve. Theory deviates appreciably from experiment under this condition.

§ 5. Solubility of He³ in liquid He⁴

The distribution coefficients, C_V/C_L, are calculated
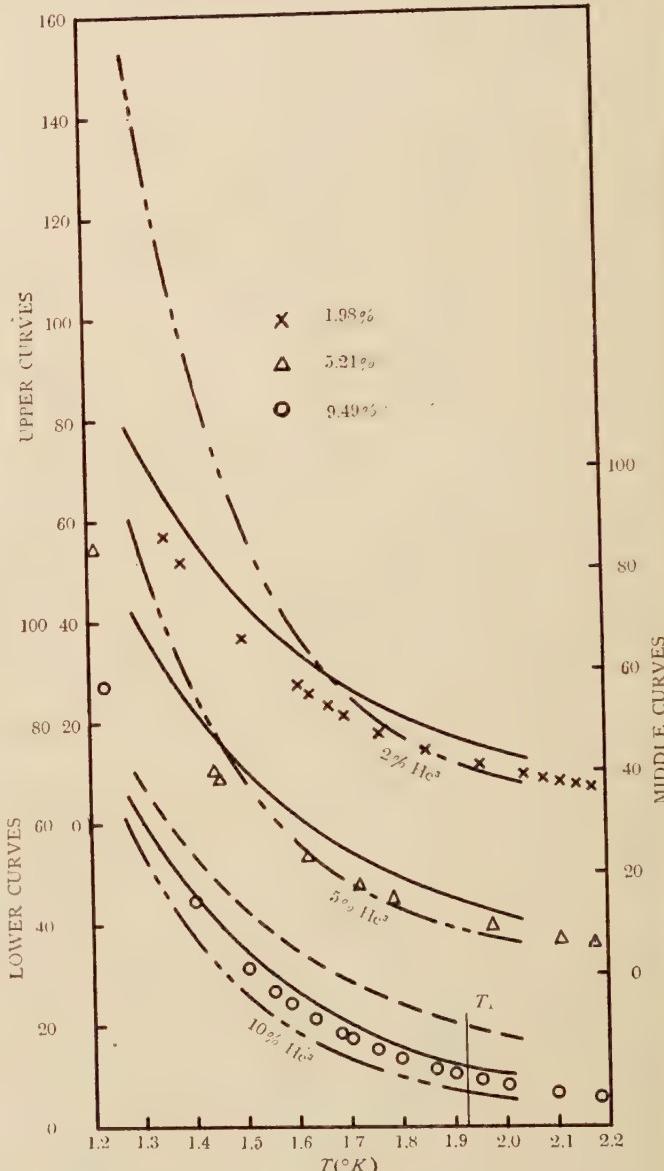
Fig. 8. Relation between the distribution coefficient, C_V/C_L, and temperature.

———— : "Modified Bose-Einstein liquid" theory : $\varDelta \propto N_4/V$.

– – – – : "Modified Bose-Einstein liquid" theory : $\varDelta =$ const.

—··— : de Boer and Gorter's theory.

Experimental points : Sommers' unsmoothed data.[16]

based on both de Boer and Gorter's and the "modified B. E. liquid" theories and the results are compared with Sommers'[16] experimental data. The concentration in the vapour phase, C_V, and the concentration in the liquid phase, C_L, are given by

$$C_V = p_3/p_4, \quad C_L = N_3/N_4 = \xi_3/1 - \xi_3$$

respectively. Thus we get

$$C_V/C_L = p_3/p_4 \Big| N_3/N_4. \qquad (45)$$

Fig. 8 shows the results of computation for C_V/C_L using (45). The experimental points included in the figures are Sommers' unsmoothed data. The discrepancies of de Boer and Gorter's theory from experiment, particularly at low concentrations and at low temperatures, are more clearly shown in this case than in the case of vapour pressure. The dashed curve in Fig. 8 corresponds to the assumption $(\partial \varDelta/\partial N_3)_{N_4} = (\partial \varDelta/\partial N_4)_{N_3} = 0$. The importance of the dependence of $\varDelta$ on concentration is again clearly seen.

That the two theories seem to deviate the more from each other the lower becomes the concentration, tempts us to compare them in the limiting case of low concentration. We have constructed at $T = 1.3°$K the $p - \xi_3$ diagram in the low concentration range, as
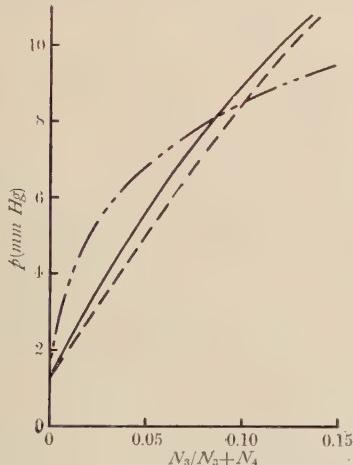
Fig. 9. Vapour pressure *versus* He³-concentration
relation in the low concentration range.
————: "Modified Bose-Einstein liquid" theory.
—·—·—: de Boer and Gorter's theory.
— — — : Sommers' experiment.
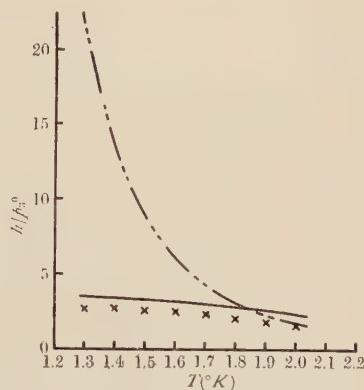
Fig. 10. Henry's law coefficient divided by p_3^0 as
a function of temperature.
————: "Modified Bose-Einstein liquid" theory.
—·—·—: de Boer and Gorter's theory.
×   : Derived from Sommers' smoothed vapour
pressure data for a 1.0 percent solution.

shown in Fig. 9. As is clearly seen in the figure, the "modified B. E. liquid" theory agrees fairly well with experiment, particularly in confirming Henry's law to hold up to the concentrations of several percent, while according to de Boer and Gorter's theory the deviation from Henry's law should already begin at much lower concentrations. The Henry's law coefficient, h, if defined by the equation*

---

* $h = p_3^0$ when Raoult's law is true.

$$h= \lim_{\xi_3 \to 0} p_3/\xi_3,$$

should be given, according to the " modified B. E. liquid " theory, by

$$h=p_3^0 \frac{V_3^0}{V_4^0} \exp\left[1-\frac{1}{\nu} \frac{V_3^0}{V_4^0}\left(1+\frac{\varDelta_0}{T}\right)x_0\right]. \tag{46}$$

According to de Boer and Gorter, it should be

$$h=p_3^0/x_0. \tag{47}$$

Fig. 10 shows the values of h/p_3^0 calculated from (46) or (47), as a function of temperature. The observed points plotted in the same figure are calculated from Sommers' vapour pressure data concerning 1.0 percent solution, applying the relation

$$p=h\xi_3+p_4^0(1-\xi_3).$$

Experiment appears to be in strong support of the " modified B. E. liquid " theory rather than of de Boer and Gorter's. This conclusion will play an important role in the later interpretation of the fact that the velocity of second sound in liquid helium remarkably increases with very small addition of He^3 (c.f. § 6)

§ 6.　The second sound

Two different versions are possible concerning the dynamical behaviour of He^3 in the case of second sound propagation in solution. The first is to suppose He^3 to stand still with the center of gravity of the liquid ($v_s=0$), and the second to suppose He^3 to move with the same velocity of the normal fraction of He^4 ($v_3=v_n$). We have calculated the velocity of second sound from both view-points. It is very interesting that the results, when compared with observation, are to support the former view-point alone if one accepts de Boer and Gorter's theory, and the latter alone if one accepts the " modified B. E. liquid " theory. The first part of this conclusion has been already announced by Koide and Usui[5], who claimed therefrom that He^3 must not partake in the normal part motion. But, since this conclusion, on the other hand, seems to be inconsistent with those observations, in which He^3 partakes in normal part of particles motion accompanied by a stationary heat flow, it will be rather welcome that the " modified B. E. liquid " theory leads to the conclusion that He^3 moves with the same velocity of the normal part of He^4 in second sound.

Case (i). $v_3=0$. The equations of motion for the superfluid and the normal parts, in this case, are :

$$(1-x_3)\partial v_s/\partial t=-1/\rho \cdot (1-x_3)\mathrm{grad}\, p+xG_{xx}\,\mathrm{grad}\, x, \tag{48}$$

$$(1-x_3)\partial v_n/\partial t=-1/\rho \cdot (1-x_3)\mathrm{grad}\, p-(1-x)\,G_{xx}\,\mathrm{grad}\, x, \tag{49}$$

where $x_3=\rho_3/\rho$ and $G_{xx}=(\partial^2 G/\partial x^2)_{p,T,c_3}$, etc. Since the observed velocities of second sound may be regarded as to correspond to the limiting case of low frequency in the theory of Gorter et al[6], we use the formula

$$v_{II}^2 = \frac{x(1-x)}{1-x_3} G_{xx}\left(1 - \frac{G_{xx}G_{TT}}{G_{xT}^2}\right)^{-1}, \tag{50}$$

which is easily derived from (48) and (49) following their method.

A. *The "modified B. E. liquid" theory.* The Gibbs free energy, G, and the entropy, S, of the solution are of the form

$$G(p, T, x, x_3) = (1-x_3)\tilde{G}_4(p, T, x, x_3) + x_3 G_3(p, T, x_3), \tag{51}$$

$$S(p, T, x, x_3) = (1-x_3)\tilde{S}_4(p, T, x, x_3) + x_3 S_3(p, T, x_3), \tag{52}$$

where $\tilde{G}_4$ and $\tilde{S}_4$ refer to non-equilibrium state, in contrast to $G_4(p, T, x_3)$ and $S_4(p, T, x_3)$ defined for equilibrium state. If we assume Gorter's condition* for the part of entropy connected with the He¹-component in solution, at least near equilibrium, $\tilde{S}_4$ should have the form

$$\tilde{S}_4(p, T, x, x_3) = k/M_4 \cdot (\varDelta/T + 5/2)x \tag{53}$$

corresponding to the expressions of x and S_4 in the "modified B. E. liquid" theory, i.e.

$$x = \left(\frac{2\pi M_4 kT}{h^2}\right)^{3/2} v\frac{V}{N_4}\exp(-\varDelta/T), \tag{54}$$

$$S_4 = \frac{k}{m_4}\left(\frac{2\pi M_4 kT}{h^2}\right)\frac{V}{N_4}\left(\frac{\varDelta}{T} + \frac{5}{2}\right)\exp(-\varDelta/T). \tag{55}$$

Integrating (52) with respect to T gives

$$\tilde{G}_4(p, T, x, x_3) = k/M_4 \cdot (\varDelta \ln T + 5T/2)x + h(p, x, x_3), \tag{56}$$

where h is an unknown function not involving T explicitly. The condition that the relation

$$(\partial G/\partial x)_{p,T,x_3} = 0 = (\partial\tilde{G}_4/\partial x)_{p,T,x_3} = -k/M_4 \cdot (\varDelta \ln T + 5T/2) + (\partial h/\partial x)_{p,x_3}$$

should be equivalent to eq. (53), leads to the following relations:

$$\left.\begin{aligned}
&G_{xx} = (1-x_3)(\partial\tilde{S}_4/\partial x)_{p,T,x_3}/(\partial x/\partial T)_{p,x_3}, \\
&G_{xT} = -(1-x_3)(\partial\tilde{S}_4/\partial x)_{p,T,x_3}, \quad G_{TT} = -(1-x_3)(\partial\tilde{S}_4/\partial T)_{p,x,x_3}, \\
&G_{xx_3} = (1-x_3)(\partial\tilde{S}_4/\partial x)_{p,T,x_3}(\partial T/\partial x_3)_{p,x}.
\end{aligned}\right\} \tag{57}$$

In deriving the above expression for G_{TT} the contribution from $x_3 G_3$ in (51) is neglected, because this is small in dilute solutions and, moreover, because G_{TT} appears only in a correction term in the expression of v_{II}^2; that is, in $G_{xx}G_{TT}/G_{xT}^2$ in (49), whose absolute value is easily found to be $\lesssim 0.2$. $(\partial x/\partial T)_{p,x_3}$ and $(\partial T/\partial x_3)_{p,x}$ are considered to be

---

* Gorter's condition, $x(\partial\tilde{S}_4/\partial x) = \tilde{S}_4$, may be regarded as empirically confirmed in pure liquid He¹ at least above 1.3°K, because H. London's relation has been found to be valid there for very narrow slits. It may be not unreasonable to extend the same condition to the case of He¹ component in the mixture.

computed with the aid of the relation (54). The right-hand side of (54) depends on x_3 through the quantities

$$V/N_1 = V_4^0(1 + (V_3^0/V_4^0 - 1)\xi_3)/(1 - \xi_3), \tag{58}$$

$$\varDelta = C_1'\varDelta = \varDelta_0(1 - \xi_3)/(1 + (V_3^0/V_4^0 - 1)\xi_3), \tag{59}$$

for ξ_3 is related to x_3 by the relation $\xi_3 = 4/3 \cdot x_3/(1 + x_3/3)$.

After numerical calculations using (54), (57), (58), and (59), we find that the factor $x(G_{cc}/(1 - x_3)$ appearing on the right-hand side of (49) can be replaced by the value for pure He^4 within an error less than 0.2 percent if the solution is of the concentration less than one percent. Thus for dilute solutions ($\xi_3 \ll 0.01$), (50) becomes

$$v_{II}^2 = (1 - x)/(1 - x_0) \cdot v_{II}^{02} \tag{60}$$

with fairly good approximation, where v_{II}^0 is the velocity of second sound in pure He^4 at the temperature under consideration.

B. *de Boer and Gorter's theory.* The Gibbs free energy (23) can be transformed into

$$G(p, T, x, x_3) = (1 - x_3)G_4^0(p, T, x) + x_3 G_3^0(p, T)$$

$$- \frac{RT}{4}x(1 - x_3)\ln\left(1 + \frac{1}{x}\frac{\xi_3}{1 - \xi_3}\right) - \frac{RT}{3}x_3\ln\left[1 + \frac{x(1 - \xi_3)}{\xi_3}\right], \tag{61}$$

where G, G_4^0 and G_3^0 refer to unit mass of the solution and of pure liquids respectively. Neglecting as before the contribution of the term $x_3 G_3^0$, we get $G_{TT} = 0$ for dilute solutions, if Tisza's model is used for pure liquid He^4. Introducing (61) in (50), we obtain

$$v_{II}^2 = x(1 - x)\left(\frac{\partial^2 G_4^0}{\partial x^2}\right)_{p,T} + \frac{RT}{3}\frac{x_3}{x}\frac{(1 - x_3)(1 - x)}{1 + (4/3)(x_3/x) - x_3}. \tag{62}$$

From the expression (24) of the Gibbs free energy in Tisza's model, we find that $(\partial^2 G_4^0/\partial x^2)_{p,T}$ does not depend on T explicitly. Then we can put

$$x(1 - x)\partial^2 G_4^0/\partial x^2 = [v_{II}^0]_\tau^2 \tag{63}$$

where $[v_{II}^0]_\tau$ is the velocity of second sound in pure He^4 at such a temperature τ that the normal fraction x_0 in pure He^4 at this temperature has the same value as in the solution at the temperature T, that is, such that $[x_0]_\tau = [x]_T$. Substituting (63) into (62) gives

$$v_{II}^2 = [v_{II}^0]_\tau^2 + \frac{RT}{3}\frac{x_3}{x}\frac{(1 - x_3)(1 - x)}{1 + (4/3)(x_3/x) - x_3}. \tag{64}$$

Case (ii). $v_3 = v_n$. The equations of motion may be written as follows* :

---

* These equations may be derived by the method, following either Koide and Usui[5] or Mazur and Prigogine[17].

$$(1-x_s)\partial v_s/\partial t = -1/\rho \cdot (1-x_3)\operatorname{grad} p + [xG_{xx} + x_3(1-x_3)G_{x3x}]\operatorname{grad} x$$
$$+ [xG_{x3} + x_3(1-x_3)G_{x3x_3}]\operatorname{grad} x_3, \quad (65)$$

$$\{1-(1-x_3)(1-x)\}\partial v_n/\partial t = -1/\rho \cdot \{1-(1-x_3)(1-x)\}\operatorname{grad} p$$
$$-(1-x)[xG_{xx} + x_3(1-x_3)G_{x3x}]\operatorname{grad} x - (1-x)[xC_{x3x} + x_3(1-x_3)G_{x3x3}]\operatorname{grad} x_3.$$
$$(66)$$

Combining (65) and (66), the velocity of second sound at low frequencies is given by

$$v_{II}^2 = \frac{x(1-x)}{1-(1-x_3)(1-x)}\left(1 - \frac{G_{xx}G_{TT}}{G_{xT}^2}\right)^{-1}$$

$$\times \left[\frac{x}{1-x^3}G_{xx} + x_3\left(2 + \frac{x_3}{x}\frac{G_{xx}G_{TT}}{G_{xT}^2}\right)G_{xx} + \frac{x_3^2}{x}(1-x_3)\left(1 - \frac{G_{xx}G_{TT}}{G_{xT}^2}\right)G_{x3x3}\right]. \quad (67)$$

A. *The "modified B. E. liquid" theory.* As in Case (i), A, it is easily found that for dilute solutions ($\xi_3 < 0.01$) eq. (67) can be written with sufficient approximation as

$$v_{II}^2 = \frac{x}{1-(1-x_3)(1-x)}\left[\frac{1-x}{1-x_0}v_{II}^{02} + 2x_3(1-x)G_{x3x}\left(1 - \frac{G_{xx}G_{TT}}{G_{xT}^2}\right)^{-1}\right.$$

$$\left. + \frac{x_3^2}{x}(1-x_3)(1-x)G_{x3x3}\right]. \quad (68)$$

Using (54), (58) and (59), one easily obtains

$$\left(\frac{\partial T}{\partial x_3}\right)_{p,x} = \frac{(\varDelta + T)\dfrac{d}{dx_3}\ln\varDelta}{\dfrac{\varDelta}{T} + \dfrac{3}{2}}$$

and hence, using (57),

$$G_{x3x} = \frac{R}{4\nu}\frac{\dfrac{\varDelta}{T} + \dfrac{5}{2}}{\dfrac{\varDelta}{T} + \dfrac{3}{2}}(\varDelta + T)(1-x_3)\frac{d}{dx_3}\ln\varDelta. \quad (69)$$

For dilute solutions ($\xi_3 < 0.01$), we may put

$$d\ln\varDelta/dx_3 \approx -4/3 \cdot V_3^0/V_4^0 = -1.81.$$

Now there still remains a quantity G_{x3x3}, which should be transformed into a form convenient for practical application. This is not possible in a similar way as the various relations in (57) were derived. For dilute solutions, however, this quantity can be calculated from eq. (14) with sufficient approximation. As G here refers to unit mass of the solution, $m_3N_3 = x_3$ and $m_4N_4 = 1-x_3$. Using these relations gives

$$(\partial^2 G/\partial x_3^2)_{p,T} = 1/m_3^2 \cdot (\partial \mu_3/\partial N_3)_{p,T,N_4} - 2/m_3 m_4 \cdot (\partial \mu_3/\partial N_4)_{p,T,N_3} + 1/m_4^2 \cdot (\partial \mu_4/\partial N_4)_{p,T,N_3}.$$
$$(70)$$

Applying (14) we find that the last two terms in (70) can be neglected compared with the first term on the right-hand side, which becomes $RT/3 \cdot 1/x_3$ for dilute solutions, with good approximation. On the other hand, $G_{x_3 x_3}$ is expressed, by definition, as

$$G_{x_3 x_3} = (\partial^2 G/\partial x_3^2)_{p,T} - 2(\partial x/\partial x_3)_{p,T} G_{x_3 x} - (\partial x/\partial x_3)^2_{p,T} G_{xx}.$$

For $\xi_3 \lesssim 0.01$, the last two terms on the right-hand side can be neglected, because these terms are less than one percent of the first term in the low temperature region, and because the contribution to v''_{11} of $G_{x_3 x_3}$ itself becomes relatively small in the high tempera-ture region though the sum of the neglected terms is possible to amount to two or three percent of the total $G_{x_3 x_3}$ in this case. In consequence, we can put

$$G_{x_3 x_3} = RT/3 \cdot 1/x_3 \qquad (71)*$$

with sufficient approximation for the present purpose. Substitution of (57), (69) and (71) into (68) results

$$v''_{11} = \frac{1}{1 + \dfrac{x_3}{x} - x^3}\left[\frac{1-x}{1-x_0}v''^0_{11} - \frac{2}{3}\frac{R}{\nu}\frac{V_3^0}{V_4^0}x_3(1-x)\frac{(J+T)\left(1 + \dfrac{5T}{J} + \dfrac{25T^2}{4J^2}\right)}{1 + \dfrac{5T}{J} + \dfrac{15T^2}{4J^2}}\right.$$
$$\left. + \frac{RT}{3}(1-x_3)(1-x)\frac{x_3}{x}\right]. \qquad (72)$$

B. *de Boer and Gorter's theory.* From the definition of G, eq. (61), it follows

$$\frac{x}{1-x_3}G_{xx} + 2x_3 G_{x_3 x} + \frac{x_3^2}{x}(1-x)G_{x_3 x_3} = \frac{x}{1-x_3}\frac{\partial^2 G_4^0}{\partial x^2},$$

if one considers He⁴ in solution as a perfect classical gas with spin 1/2, as was assumed successfully in § 2. Then, putting $G_{TT}=0$ as before, we get from (67)

$$v''_{11} = \frac{1}{1 + \dfrac{x_3}{x} - x_3}\frac{x(1-x)}{1-x_3}\left(\frac{\partial^2 G_4^0}{\partial x^2}\right)_{p,T},$$

which, by substituting (63), leads to

$$v''_{11} = \frac{1}{1 + \dfrac{x_3}{x} - x_3}[v^0_{11}]^2_T. \qquad (73)$$

The observation by Lynton and Fairbank[18] of the velocity of second sound in a solution containing 0.8 percent He^3, is compared with the theoretical results corresponding to the four cases described above. Smoothed empirical values are used for v'_{II} in eqs. (60), (63), (72), and (73). The values of x_0 and x employed both in the "modified B. E. liquid" theory and in de Boer and Gorter's theory, are derived, by applying (29) and (44), from observed values of S_0, c_p^0 and v'^0_{II}. The other necessary quantities are all theoretically evaluated on the basis of the respective theories. Fig. 11 shows the results of calculation compared with experiment. It is clear that theory and experiment are in allowable agreement, either in Case (i), B($v'_3 = 0$, de Boer and Gorter) or in Case (ii), A ($v'_3 = v'_n$, "modified B. E. liquid"), but in complete disagreement in other cases. It is very interesting that the "modified B. E. liquid" theory is consistent only with the assumption that, in the case of second sound propagation, He^3 participates in
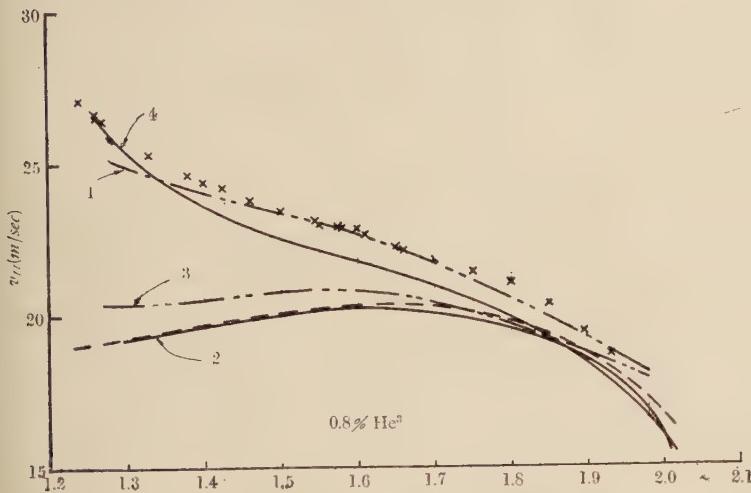
Fig. 11. Velocity of second sound. Comparison of various theories for the 0.8 percent solution. 1. $v_3 = 0$, de Boer and Gorter's theory, 2. $v_3 = 0$, "Modified B. E. liquid" theory, 3. $v_3 = v'_n$, de Boer and Gorter's theory, 4. $v_3 = v'_n$, "Modified B. E. liquid" theory. $\times$: Lynton and Fairbank's experiment.[18] — — — : Experimental curve for pure He^4.
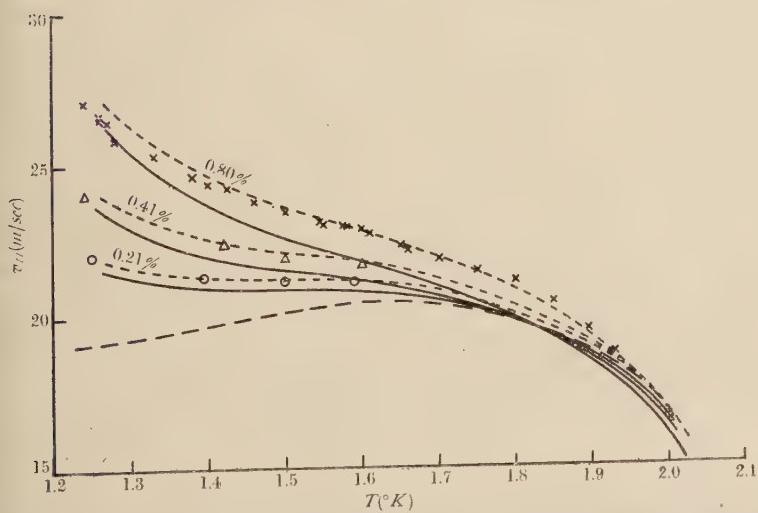
Fig. 12. Velocity of second sound in solutions of various concentrations. Comparison with experiment of the "Modified B. E. liquid" theory. ——— : According to eq. (72). — — — : According to a tentative equation obtainable from (72) by dropping the second term in [ ]. Experimental points : Lynton and Fairbank[18]. (c.f. § 10)

the motion of the normal part of He[1], while de Boer and Gorter's theory only with the alternative one that He[3] is stationary in the center of gravity system. Since de Boer and Gorter's view-point has been found to be decisively inconsistent with the observations concerning the vapour pressure in solution, one must conclude $v_3 = v_n$ in the phenomenon of second sound. The solid curves in Fig. 12 show the results of calculation as regards the 0.8, 0.41 and 0.21 percent solutions, based on the "B. E. liquid" theory under the assumption $v_3 = v_n$, together with those of observation by Lynton and Fairbank.[18]

§ 7. The osmotic pressure

It is a well-known phenomenon that the liquid level in the solution is found to be higher than the pure liquid He[4] level when the two liquids at the same temperature are connected by a very narrow slit above both liquid levels and equilibrium is established. This excess pressure in the solution of He[3] in He[4] is called "osmotic pressure"; the slit plays here the role of semi-permeable membrane in the similar phenomena well-known in usual solutions. By the experiment of Taconis et al,[19] the pressure difference $\varDelta p$, at least for very dilute solutions, was roughly found to obey the Van't Hoff equation

$$\varDelta p = RTC_3, \tag{74}$$

where C_3 is the number of moles of He[3] found in unit volume of the solution. Theoretically, eq. (74) has been derived by the consideration[20] based on de Boer and Gorter's theory, employing the expression, (23), of the Gibbs free energy for a classical ideal mixture. When de Boer and Gorter's point of view is now given up, can Van't Hoff's equation be derived in the "modified B. E. liquid" theory, too? It is the case for dilute solutions, as will be shown below.

Putting the right-hand side of eq. (65) to zero, and assuming the thermodynamical equilibrium at every point, i.e.

$$\mathrm{grad}(\partial G/\partial x)_{p,T,x_3} = 0,$$

we get

$$1/\rho \ \mathrm{grad}\ p = x(\partial S/\partial x)_{p,T,x_3}\,\mathrm{grad}\ T + x_3\,G_{x_3x}\,\mathrm{grad}\ x + x_3\,G_{xx_3}\,\mathrm{grad}\ x_3. \tag{75}$$

When both liquids have the same temperature and one of them is pure He[4], (75) integrates to

$$\varDelta p = RTC_3\left[x_3 C_4'\left\{\frac{1}{C_3'} + 1 - \frac{1}{\nu}\frac{V_3^0}{V_4^0}\left(\frac{2\varDelta}{T} + 1\right)x\right\}\frac{d}{dx_3}\ln\varDelta\right.$$

$$\left. - \frac{3}{8\nu}x_3 \colon x\left(\frac{\varDelta}{T} + \frac{5}{2}\right)^2\frac{1 + \frac{5T}{2\varDelta}}{1 + \frac{3T}{2\varDelta}}\left(\frac{d}{dx_3}\ln\varDelta\right)^2\right], \tag{76}$$

where the quantities under the sign (- - - -) are to be averaged appropriately between the concentrations $(0,\ x_3)$. For dilute solutions, this average may be replaced by the value

at the concentration $x_3/2$. In deriving (76), G_{x_3x} and $G_{x_3x_3}$ are calculated from (14) by differentiating it partially with respect to x and x_3 respectively, the difference between $1/m_3 \cdot (\partial \mu_3/\partial x)_{p,T,x_3}$ and G_{x_3x}, or between $1/m_3 \cdot (\partial \mu_3/\partial x_3)_{p,T,x}$ and $G_{x_3x_3}$, being neglected as we did in § 6 when we derived (71). Let the right-hand side of (76) be denoted by $RTC_3(\alpha - \beta)$. The deviation of the factor, $\alpha - \beta$, from unity means the departure from Van't Hoff's law. Table I shows the values of $\alpha - \beta$ for the two solutions with the concentrations of 2.0 and 5.0 percent, together with the values of α. The latter corresponds to the correction factor, for Van't Hoff's equation, to be expected in the case of the term in G_{x_3x} being omitted from eq. (75). The table shows that deviations from Van't Hoff's law are negligibly small in those dilute solutions ($\xi_3 < 0.01$) that are usually employed for the observation of the osmotic pressure in liquid helium.

Table I. Correction factor for Van't Hoff's equation.

T	0.02		0.05	
(°K)	$\alpha - \beta$	α	$\alpha - \beta$	α
1.3	0.98	0.98	0.94	0.96
1.5	0.97	0.98	0.93	0.97
1.7	0.96	0.99	0.91	0.97
1.9	0.96	0.99	0.89	0.99

§ 8.   The thermomechanical effect

In this section, the thermomechanical effect in solutions will be discussed from the point of view based on the " modified B. E. liquid " theory. In a similar way as we have derived (76), we can easily obtain from (75), by putting grad $x_3 = 0$,

$$\Delta p = \rho S_4 \left\{ 1 + \left(\frac{\partial x}{\partial T} \right)_{p,x_3} \frac{x_3}{x} T \frac{1 + \dfrac{T}{\Delta}}{1 + \dfrac{3T}{2\Delta}} \frac{d}{dx_3} \ln \Delta \right\} \Delta T$$

$$= \overline{\rho S_4 \left\{ 1 + x_3 \left(\frac{\Delta}{T} + 1 \right) \frac{d}{dx_3} \ln \Delta \right\}} \Delta T, \tag{77}$$

where the quantity under the sign (——) should be averaged appropriately from T to $T + \Delta T$. When ΔT is small, this average may be replaced by the value at $T + \Delta T/2$. Setting $x_3 = 0$ in (77) one obtains, of course, the H. London equation.[21] As the density ρ is nearly independent of temperature, the difference ΔH of the two liquid levels is given by

$$\Delta H = 1/g \cdot \overline{S_4 \{ 1 + x_3 (\Delta/T + 1) \, d \ln \Delta/dx_3 \}} \tag{78}$$

where g is the acceleration for gravity. The calculated values of the ratio $\Delta H/\Delta H_0$, ΔH_0 referring to the case of pure liquid He⁴, are listed in Table II in the limit $\Delta T \to 0$. It is a remarkable fact that the values of $\Delta H/\Delta H_0$ are less than unity, though the entropies of solutions should be larger than those of pure He⁴. The situation becomes clear if we tentatively omit the second term in the waved brackets in (78). The values of $\Delta H/\Delta H_0$ in this case are also included in Table II and denoted by $[\Delta H/\Delta H_0]$. The comparison

of $\Delta H/\Delta H_0$ and $[\Delta H/\Delta H_0]$ exhibits the importance of the role in the thermomechanical effect of the term involving G_{x_3c} in (75), in contrast to the case of the osmotic pressure. It is due to the existence of this term that ΔH in solution becomes less than ΔH_0.

Table II.   The ratio of the thermomechanical level difference in solution to that in pure He[4].

T	$\xi_3=0.02$		$\xi_3=0.05$		$\xi_3=0.07$	
(°K)	$\Delta H/\Delta H_0$	$[\Delta H/\Delta H_0]$	$\Delta H/\Delta H_0$	$[\Delta H/\Delta H_0]$	$\Delta H/\Delta H_0$	$[\Delta H/\Delta H_0]$
1.3	0.96	1.20	0.80	1.59	0.76	1.90
1.5	0.96	1.18	0.84	1.50	0.82	1.76
1.7	0.97	1.16	0.86	1.44	0.86	1.65
1.9	0.97	1.14	0.88	1.39	0.89	1.58

§ 9.   The specific heat and $(d\dot{Q}/dR)_{\text{ini.}}$ in the analogue, in solution, of Mendelssohn and Chandrasekhar's experiment[*22)]

The experiments on the vapour pressures and the related phenomena seem to support the " modified B. E. liquid " theory of He³-He⁴ mixtures rather than de Boer and Gorter's theory.   It is desirable, however, that further experiments will be performed in future which may be of value to help decide which is the proper theory.   The specific heat and $(d\dot{Q}/dR)_{\text{ini.}}$ in the analogue, in solution, of Mendelssohn and Chandrasekhar's experiment are chosen and discussed in this section as examples.

According to the " modified B. E. liquid " theory, the specific heat is easily shown to be

$$C_v=(1-x_3)\frac{k\Delta^2}{\rho_1}\left(\frac{2\pi M_1 k}{h_2}\right)^{3/2}T^{-1/2}\left(1+\frac{3T}{\Delta}+\frac{15T^2}{4\Delta^2}\right)\exp\left(-\Delta/T\right)+C_v^3, \qquad (79)$$

where C_v^3 is the direct contribution to specific heat of the He³-component.   The specific heat of the solution, according to de Boer and Gorter's theory, is calculated from

$$C_v=(1-x_3)T\left(\frac{\partial x}{\partial T}\right)_{x_3}\left\{S_\lambda-\frac{R}{4}\ln\left(1+\frac{\xi_3}{1-\xi_3}\frac{1}{x}\right)\right\}+C_v^3, \qquad (80)$$

where $S_\lambda=0.405$ cal/g.deg. and $(\partial x/\partial T)_{c_3}$ should be computed numerically by using the equation

$$2.186\, S_\lambda x^{1/5.5}-TS_\lambda=\frac{RT}{4}\ln\left(1+\frac{\xi_3}{1-\xi_3}\frac{1}{x}\right). \qquad (81)$$

If He³ is considered as a perfect classical gas, then

$$C_v^3=RT/2\cdot x_3, \qquad (82)$$

----

* $\dot{Q}$ is the heat input required to maintain the temperature constant when He⁴ with no entropy is flowing into the solution with the rate R. $(d\dot{Q}/dR)_{\text{ini.}}=\lim_{R\to0}(d\dot{Q}/dR)$.

which is the upper limit theory may expect. Fig. 13 shows the theoretical specific heat calculated from (79), (80) and (82) as well as that obtained when C_v^3 is neglected. The areas between the two curves, shaded parts in the figures, represent the uncertainty due to the lack of our knowledge regarding the behaviour of He³ in solution. At lower temperatures de Boer and Gorter's theory predicts considerably higher specific heat of solution than the "modified B. E. liquid" theory does.

Fig. 13. Increase of the specific heats of solutions. (theoretical curves). Shaded areas represent the uncertainty due to the unknown direct contribution of He³-component.

The two theories, however, exhibit more pronounced difference in the prediction of the values of $(d\dot{Q}/dR)_{\text{ini}}$ in an experiment, in solution, of the Mendelssohn and Chandrasekhar type. Moreover, the ambiguity due to the direct contribution of the He³-component is absent in this case. Mendelssohn and Chandrasekhar made the following experiment with pure He⁴. A small Dewar vessel is half immersed in the helium bath. The Dewar vessel and the bath are connected with helium film through a very narrow slit. When heat is supplied by an electric heater to the liquid in the Dewar vessel, liquid helium flows in through the film. The heating rate, $\dot{Q}$, and the rate of flow, R, should be connected by the equation

$$\dot{Q} = \rho_4 S_4 T R \tag{83}$$

for small heat input. Experiment confirmed this proportionality, for small heat input, between $\dot{Q}$ and R, and the initial slope $(d\dot{Q}/dR)_{\text{ini}}$ gave the specific heat of pure He⁴ in good agreement with Keesom's direct measurement. Suppose now a similar experiment to be performed with the solution. Then simple calculations show that according to de Boer and Gorter's theory

$$\left(\frac{d\dot{Q}}{dR}\right)_{\text{ini.}} = \rho_4 T \left\{ S_\lambda + \frac{R}{4} x \ln\left(1 + \frac{\xi_3}{1-\xi_3} \frac{1}{x}\right) \right\} \left\{ 1 - \frac{x_3}{x}(1-x_3)\left(\frac{\partial x}{\partial x_3}\right)_{p,T} \right\}, \tag{84}$$

and according to the "modified B. E. liquid" theory

$$\left(\frac{d\dot{Q}}{dR}\right)_{\text{ini.}} = \rho_4 T S_4 \left[1 - x_3(1-x_3)\left\{ \frac{1}{x}\left(\frac{\partial x}{\partial x_3}\right)_{p,T} + \frac{d\Delta/dx_3}{\Delta + 5T/2} \right\} \right]. \tag{85}$$

$(\partial x/\partial x_3)_{p,T}$ in eq. (84) should be calculated from (81), and $(\partial x/\partial x_3)_{p,T}$ in eq. (85) is obtained from (55), (58), and (59) as

$$(\partial x/\partial x_3)_{p,T} = -x(\Delta/T + 1)\, d \ln \Delta/dx_3.$$

Substituting this into (85), one obtains

$$(d\dot{Q}/dR)_{\text{ini.}} = \rho_4 T S_4 \{1 + x_3(1-x_3)$$
$$\times (\varDelta/T + 1 - 1/(1 + 5T/2\varDelta))d\ln\varDelta/dx_3\}$$
$$\tag{86}$$

in the " modified B. E. liquid " theory.

It is to be noted that in deriving these equations the following assumptions are made (i) the liquid flowing in through the film consists of He^4 atoms alone, and (ii) the temperature and the concentrations of He^4, He_n^4, and He^3 are all kept uniform everywhere in the liquid. The results of computations by applying (84) and (85) for the 0.5 and 1.0 percent solutions are shown in Fig. 14. The difference between the predictions by both theories is more appreciable in this case than in the case of specific heat. It is remarkable that, according to the " modified B. E. liquid " theory, the value of $(d\dot{Q}/dR)_{\text{ini.}}$ is nearly independent of He^3-concentration for dilute solutions $(\xi_3 < 0.01)$*, in contrast to its appreciable increase even for such solutions in de Boer and Gorter's theory.
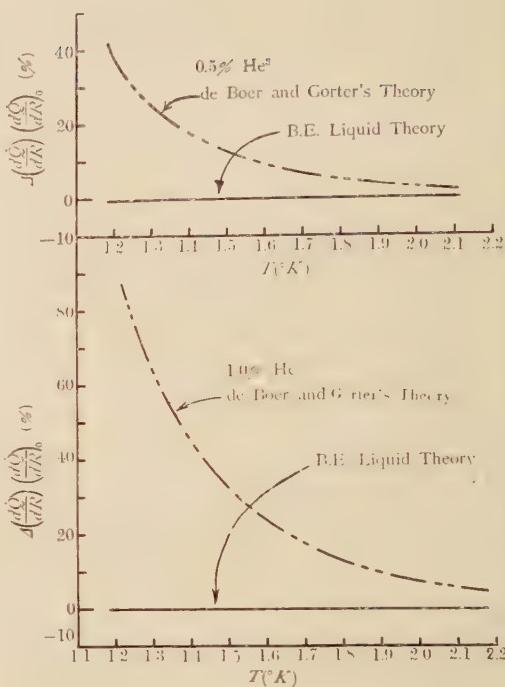
Fig. 14.   Increase of $(d\dot{Q}/dR)_{\text{ini.}}$ expected in a similar experiment for solution, as Mendelssohn and Chandrasekhar[22] carried out with pure He^4. ——— : " Modified B. E. liquid " theory, ·· ·· : de Boer and Gorter's theory.

§ 10.   Discussion

1.   *On the vapour pressure*

It is to be noted here that the degeneracy temperature for pure liquid He^3 should be $4.85°K$ if one assumes it as a perfect Fermi gas with the liquid density. If this supposition were true, pure liquid He^3 would be degenerate in the temperature range under consideration. This is, however, in conflict with the fact connected to vapour pressure, at least from the point of view based on the " modified B. E. liquid " theory, since Heer and Daunt's consideration[2] retains its validity also in the present theory : according to these authors, the actual degeneracy temperature in pure He^3 is supposed to lie below $2°K$. This lowering of degeneracy temperature in actual liquid He^3 might be ascribed to the possible increase of the particle mass and decrease in the number density in the liquid

---

* With increasing concentration, $(d\dot{Q}/dR)_{\text{ini.}}$ in the " modified B. E. liquid " theory decreases, though very slowly, leading to the value 4 percent less than that for pure He^4 at 3 percent concentration and $1,3°K$,

state, quite similarly as it was really necessary to assume for liquid He^4. But this seems not to be the case, as will be shown by the following consideration. If we assume for the He^3-particle in the liquid to have a mass different from m_0, i.e. $M_3 = \nu_3 m_0$, then (34) and (35) become

$$p_3 = p_3^0 C_3'^{1/\nu_3} \exp\left[C_4'\left\{\frac{1}{\nu_3} - \frac{1}{\nu}\frac{V_3^0}{V_4^0}\left(1 + \frac{\varDelta}{T}\right)x\right\}\right],$$

$$p_4 = p_4^0 \exp\left[-\frac{V_4^0}{V_3^0}\frac{C_3'}{\nu_3} - \frac{C_4'}{\nu}\left(x - \frac{x_0}{C_4} - \frac{N_3}{N_4}\frac{V_3^0}{V_4^0}\frac{\varDelta}{T}x\right)\right].$$

p_3 for dilute solutions is approximately proportional to ξ_3^{1/ν_3}, and if ν_3 were really different from unity, the departure from Henry's law would be observed, contrary to the fact. Thus we must conclude that the He^3-particle in the liquid state has the same mass as the He^3-atom, contrary to the case of liquid He^4.

2. *On the second sound above* 1.3°K

The calculated velocities of second sound according to the " modified B. E. liquid " theory are not still in complete accord with experiment. (c.f. Fig. 12) The possible errors in the determination of the He^3-concentration on the experimental side cannot explain this discrepancies. For, if one assumes a larger concentration in the theory, the solid curve corresponding to 0.8 percent solution in Fig. 12, for example, shifts upwards at temperatures below about 1.8°K but downwards at higher temperatures, and vice versa if a smaller concentration is assumed. Thus the agreement would be improved in neither case. The discrepencies should be then attributed to deficiency in the theory. The determination of $\varDelta$ in solution is the most questionable procedure in the theory. But it is unlikely that the disagreements are mainly caused from a fault in the estimation of x, because, even if use were made of such a small value of $[x_0]_{obs.}$, they would not entirely be removed. It would be worth noting that, if we tentatively omit the term in G_{x3x} from (68), the agreement between theory and experiment becomes almost satisfactory, as is shown by the dotted curves in Fig. 12. This fact tempts us to suppose, either that the actual value of $d\varDelta/dx_3$ is much smaller than expected from eq. (27), or that the terms in G_{x3x} appearing in (65) and (66) should not appear in the correct equations of motion. It has already been pointed out and stressed that the first assumption cannot be accepted in the problems of transition temperature and vapour pressure (§ 3, § 4, and § 5). There is still the possibility, however, that the change of $\varDelta$ cannot immediately follow such a rapid change of the concentration of He^3 as occurs in the second sound wave. Whether this is really true or not will not be clear till the molecular theory reveals the nature of the excitation energy.

The problem whether the terms proportional to G_{x3x} exist in equations of motion seems to be solved most easily by experiment. As was seen in the previous sections, while the osmotic pressures are indifferent to the omission of these terms from the equations of motion (c.f. Table I), the predicted ratio of the thermomechanical level difference in solution to that in pure He^4 increases appreciably if those terms are removed (c.f. Table II),

It is very desirable to examine the validity of the equations of motion, such as (65) and (66), by making experiments on the thermomechanical effect in solution.

3. *On the second sound near the absolute zero of temperature.*

We assume that the two fluid version for liquid helium is at least approximately valid also in this temperature region. Although neither the "modified B. E. liquid" model nor the Tisza model can be used for calculating the normal fraction x in this temperature region, x undoubtedly tends to zero very rapidly when the absolute zero of temperature is approached. Thus the second sound can be interpreted as a propagation in a wave form of the relative motion between the superfluid consisting of He^4-atoms alone and the normal fluid consisting nearly exclusively of He^3-atoms. Only the two cases in § 6, *Case* (i), B and *Case* (ii), A, which are in reasonably good agreement with experiment at high temperatures, are considered.

Assuming $x_3/x \gg 1$, eq. (64) for Case (i), B leads to

$$v_{II}^2 = [v_{II}^0]_\tau^2 + RT(1-x_3)/4. \tag{87}$$

Thus, in this theory, the velocity of second sound in dilute solutions tends to the same finite limiting value with decreasing temperature as that in pure He^4.

The first term in [ ] of eq. (72) for Case (ii), A need not be altered and the second term may be omitted, because the thermodynamic properties of liquid helium in this temperature region are principally determined by phonons but little connected with the excitations with the excitation energy Δ, and because the properties of phonons will not be appreciably affected by the addition of a small amount of He^3. The last term remains unaltered, because relation (71) merely represents the diluting effect of He^4 component but has nothing to do with its intrinsic property. Then, for x_3, $x \ll 1$ and $x_3/x \gg 1$, we get from (72)

$$v_{II}^2 = \frac{1}{1+(x_3/x)}\left[v_{II}^{02} + \frac{RT}{3}(1-x_3)\frac{x_3}{x}\right]. \tag{88}$$

Since v_{II}^{02} is approximately equal to a constant value $1.52^2 \times 10^5$ cm^2/sec^2 for $T \lesssim 0.5°$K, eq. (88) is roughly replaced, for small concentrations, by

$$v_{II}^2 = RT/3 \tag{89}$$

if x_3/x in this temperature region can be assumed to be larger than say 10^3. Formula (89) leads to the result that v_{II} tends to zero with $\sqrt{T}$ as T approaches to zero and the initial slope of the v_{II}^2-T curve should be $R/3$. This result is diametrically divergent from the above mentioned result of de Boer and Gorter's case assuming $x_3=0$.

Fig. 15 shows the results of calculation according to (89) (the solid curve starting at the origin) compared with the recent experiment by King and Fairbank[23]. The dotted curves show the results for the 0.3 and 4.3 percent solutions, calculated by using eq. (88). The value of x employed for this calculation, is derived by provisionally appealing to Landau's formula

$$x = 4/3 \cdot U_{ph}/v_I^2 \tag{90}$$

where v_1 is the velocity of first sound which is taken as a constant value, 237 m/sec., and U_{ph} is the energy of phonon which is here replaced by $0.0235/4 \cdot T^4$ joule/g, the total energy observed by Kramers *et al*[13]. It is clearly seen from the figure that the conditions for the approximate validity of (89) are nearly fulfilled for these He^3-concentrations and the value of x given by (90), and that the prediction of (88) or (89) is qualitatively confirmed by experiment.* It is to be noted that (87) always gives the values of v_{II} larger than $v_{II}^0 = 152$ m/sec., contrary to the fact. It is very desirable to extend the experiment further to lower temperatures and to investigate whether v_{II}^0 is really proportional to T and, if so, whether the proportionality coefficient is given by $R/3$ as stated above. The same result seems to have been reached by Pomeranchuk[21]** from somewhat different considerations. The observed absence of dispersion in second-sound velocity for dilute solutions at low temperatures can be accounted for by saying that the second sound in solution is there essentially of the same nature as that observed for pure He^4 at high temperatures, that is, the second sound in both cases is interpreted as the propagation in a wave form of relative motion between two kinds of fluid, in contrast to the propagation of the fluctuation of thermal energy as supposed to be the case in pure He^4 at low temperatures.[25]

Fig. 15. Velocity of second sound near absolute zero of temperature. Solid curve: eq. (89). Upper dotted curve: eq. (88), the 0.3 percent solution. Lower dotted curve: eq. (88), the 4.3 percent solution. Experimental points: King and Fairbank.[23]

In conclusion, the author wishes to thank his colleagues, Messrs. A. Morita, N. Honda, C. Horie, and Y. Fukuda, for their valuable advices.

* If we assume a mass $M_3 = \nu_3 m_3$ for the particle in liquid He^3, we obtain $v_{II}^0 = 1/\nu_3 \cdot RT/3$. To take ν_3 considerably larger than unity, similarly as in the case of liquid He^4, would again contradict the experiment. It is an interesting problem how statistics concerns this striking contrast between the particle masses for liquid He^3 and liquid He^4.

** Details are not clear to me as his paper is not available here.

References

1) J. de Boer and C. J. Gorter, Physica **16** (1950), 225.
2) C. V. Heer and J. G. Daunt, Phys. Rev. **81** (1951), 447.
3) A. Bijl, J. de Boer, and A. Michels, Physica **8** (1941), 655.
4) M. Toda, Prog. Theor. Phys. **6** (1951), 458.
5) S. Koide and T. Usui, Prog. Theor. Phys. **6** (1951), 506.
6) F. London, Phys. Rev. **54** (1938), 947.
7) C. J. Gorter, P. W. Kasteleijn, and J. H. Mellink, Physica **16** (1950), 113.
8) C. J. Gorter, Physica **15** (1949), 523.
9) A. C. Hollis-Hallett, Proc. Roy. Soc. **210** (1951), 404.
10) E. F. Hammel and A. F. Schuch, Phys. Rev. **87** (1952), 154.
11) E. J. Burge and L. C. Jackson, Proc. Roy. Soc. **205** (1951), 270.
12) K. Mendelssohn and G. K, White, Proc. Phys. Soc, **63A** (1950), 1328.
13) H. A. Kramers, J. D. Wasscher, and C. J. Gorter, Physica **18** (1952), 329
14) J. R. Pellam, Phys. Rev. **75** (1949), 1183.
15) R. D. Maurer and M. A. Herlin, Phys. Rev. **76** (1949), 869.
16) H. S. Sommers, Phys. Rev. **88** (1952), 113.
17) P. Mazur and I. Prigogine, Physica **17** (1951), 680.
18) E. A. Lynton and H. A. Fairbank, Phys. Rev. **80** (1950), 1043.
19) K. W. Taconis, J. J. M. Beenakker and Z. Dokoupil, Phys. Rev. **73** (1948), 638.
20) C. J. Gorter, K. W. Taconis, and J. J. M. Beenakker, Physica **17** (1951), 841.
21) H. London, Proc. Roy. Soc. **171** (1939), 484.
22) K, Mendelssohn and B. S. Chandrasekhar, Proc. Oxford Conf. (1951), 67.
23) J. C. King and H. A. Fairbank, Phys. Rev. **90** (1953), 989.
24) I. Pomeranchuk, J. Exp. Theor. USSR. **19** (1949), 42.
25) J. C. Ward and J. Wilks, Phil. Mag. **42** (1951), 314 ; Phil. Mag. **43** (1952), 48.

On the Spontaneous Magnetizations of Honeycomb and Kagomé Ising Lattices

Shigeo NAYA

Department of Physics, Osaka University

The spontaneous magnetizations for the honeycomb and kagomé lattice are obtained by transforming Potts' results for the triangular lattice. Potts' relations are discussed.

§ 1. Introduction

Although the statistical theories of the two dimensional Ising ferromagnet have been treated exactly by several authors for the various types of the lattices, the similar problem in the presence of an external field have not yet been attacked in an exact way. For an evaluation of spontaneous magnetizations, however, we can utilize the method of perturbation calculation in solving the eigenvalue problem in the limiting case of vanishing external field: $\mathfrak{H} \to 0$. In fact, C.N. Yang has performed the perturbation calculation and evaluated exactly the spontaneous magnetization for the two dimensional square lattice. His result is surprisingly simple, in spite of the complicated nature of the other thermodynamical quantities. Potts has generalized Yang's result to the case of the triangular lattice with an equally simple result. We shall try to find the spontaneous magnetizations for the honeycomb and kagomé lattices by transforming Potts' result for the triangular lattice.

§ 2. Semi-ferromagnetic honeycomb lattice

Let us consider a lattice to be explained below. (see Fig. 1) In the honeycomb lattice divided into the two triangular sublattices, all the lattice points which belong to one triangular sublattice (denoted by ●) have the magnetic moment: m_0 and all the lattice points which belong to the other sublattice (denoted by ○) have no magnetic moment. In the presence of an external field, the magnetic energy of this honeycomb lattice will be

$$\mathfrak{H}(\mu_1 + \mu_3 + \mu_5 + \cdots) + \mathfrak{H}(\mu_2' + \mu_4' + \cdots) + \mathfrak{H}(\mu_1'' + \mu_3'' + \cdots) + \cdots . \tag{1}$$

$$\mathfrak{H} = -m_0 H / kT , \tag{2}$$

where μ_i's are the Ising spin variables which take on either of the two values 1 or -1. On the other hand, we assume that the interaction energy for a pair of neighboring atoms
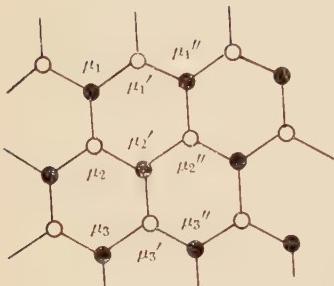
Fig. 1

of this lattice is the same as that of the ordinary Ising honeycomb lattice. Thus only the half of atoms contribute to its magnetic properties (especially to its spontaneous magnetization). Let us call such a type of the lattice "Semi-ferromagnetic lattice". On the other hand, let us call "Ferromagnetic lattice" the ordinary lattices where all the atoms have the magnetic moment m_0. To relate the spontaneous magnetization of the ferromagnetic honeycomb lattice to that of the semi-ferromagnetic honeycomb lattice is our object in this section.

The operator V of the eigenvalue problem for the semi-ferromagnetic honeycomb lattice is given by (Fig. 2)

$$V = V_{H,1} \cdot V_1 \cdot V_{H,2} \cdot V_2, \quad (V_0 = V_1 \cdot V_2) \tag{3}$$

$$V_{H,1} = e^{\mathfrak{H}(s_1 + s_3 + s_5 + \cdots)}, \quad V_{H,2} = e^{\mathfrak{H}(s_2 + s_4 + s_6 + \cdots)},$$

$$V_1 = e^{L(s_1 s_2 + s_3 s_4 + \cdots)} \cdot e^{L^*(C_1 + C_2 + C_3 + \cdots)},$$

$$V_2 = e^{L(s_2 s_3 + s_4 s_5 + \cdots)} \cdot e^{L^*(C_1 + C_2 + C_3 + \cdots)}, \tag{4}$$

where we have dropped a scalar factor. The eigenvalue equation is

$$V\Psi = \lambda\Psi. \tag{5}$$
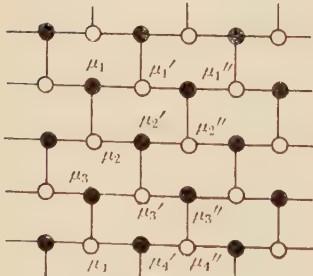
Fig. 2

For evaluation of the spontaneous magnetization, wa can adapt the method of the perturbation calculation in the limiting case of vanishing external field : $H \rightarrow 0$.

$$V = V_0 + \mathfrak{H} V_1 + \cdots,$$

$$\lambda = \lambda_0 + \mathfrak{H}\lambda_1 + \cdots \tag{6}$$

where $V_0 = V_1 \cdot V_2$ is the operator for the eigenvalue problem in the case of the vanishing field and λ_0 is the maximum eigenvalue of V_0 which is known to be doubly degenerate below the Curie temperature.

$$V_0\Psi_0^+ = \lambda_0\Psi_0^+,$$

$$V_0\Psi_0^- = \lambda_0\Psi_0^-. \tag{7}$$

Now the partition function for the semi-ferromagnetic honeycomb lattice in the first order perturbation is

$$f' = \lambda^{2m} = (\lambda_0 + \mathfrak{H}\lambda_1 + \cdots)^{2m} = \lambda_0^{2m} + 2m\mathfrak{H}\lambda_0^{2m-1} \cdot \lambda_1 + \cdots. \tag{8}$$

The total spontaneous magnetization I' of the semi-ferromagnetic honeycomb lattice is given by

$$I' = \partial \log f'/\partial \mathfrak{H} \mid_{\mathfrak{H} \to 0} = 2m\lambda_0^{2m-1} \cdot \lambda_1/\lambda_0^{2m} = 2m \cdot \lambda_1/\lambda_0, \tag{9}$$

where λ_1 is determined from the secular equation of the perturbation. In order to estimate λ_1, let us expand the operator V in the power series of the external field: $\mathfrak{H}$.

$$
\begin{aligned}
V &= V_1 \cdot V_2 + \mathfrak{H}(s_1 + s_3 + s_5 + \cdots) V_1 V_2 + \mathfrak{H} V_1 (s_2 + s_4 + s_6 + \cdots) V_2 + \cdots \\
&= V_1 \cdot V_2 + \mathfrak{H}\{(s_1 + s_3 + s_5 + \cdots) + V_1 (s_2 + s_4 + s_6 + \cdots) V_1^{-1}\} V_1 V_2 + O(\mathfrak{H}^2) \\
&= V_1 V_2 + \mathfrak{H} V_1 + O(\mathfrak{H}^2). \tag{10}
\end{aligned}
$$

V_1 is nondiagonal because it anticommutes with $U = c_1 c_2 c_3 \cdots c_{2n}$. It is easy to find the following relations as proved by Yang.

$$(\varPsi_0^{+\prime} V_1 \varPsi_0^+) = 0, \quad (\varPsi_0^{-\prime} V_1 \varPsi_0^-) = 0.$$

$$
\begin{aligned}
(\varPsi_0^{+\prime} V_1 \varPsi_0^-) &= (\varPsi_0^{-\prime} V_1 \varPsi_0^+) = \{(\varPsi_0^{+\prime}(s_1 + s_3 + s_5 + \cdots)\varPsi_0^-) \\
&+ (\varPsi_0^{+\prime} V_1 (s_2 + s_4 + s_6 + \cdots) V_1^{-1} \varPsi_0^-)\} \lambda_0 \equiv \{\varDelta_1 + \bar{\varDelta}_2\}\lambda_0. \tag{11}
\end{aligned}
$$

Therefore the secular equation becomes as follows:

$$
\begin{vmatrix}
(\varPsi_0^{+\prime} V_1 \varPsi_0^+) - \lambda_1, & (\varPsi_0^{+\prime} V_1 \varPsi_0^-) \\
(\varPsi_0^{-\prime} V_1 \varPsi_0^+), & (\varPsi_0^{-\prime} V_1 \varPsi_0^-) - \lambda_1
\end{vmatrix}
=
\begin{vmatrix}
-\lambda_1, & (\varDelta_1 + \bar{\varDelta}_2)\lambda_0 \\
(\varDelta_1 + \bar{\varDelta}_2)\lambda_0, & -\lambda_1
\end{vmatrix}
= 0, \tag{12}
$$

$$\lambda_1 = \pm (\varDelta_1 + \bar{\varDelta}_2) \cdot \lambda_0.$$

Thus the degeneracy of the maximum eigenvalue of V_0 is removed by the external field and it is divided into the two levels $\lambda = \lambda_0 + \mathfrak{H} \mid \lambda_1 \mid$, $\lambda = \lambda_0 - \mathfrak{H} \mid \lambda_1 \mid$. Of course, we must choose the maximum eigenvalue $\lambda = \lambda_0 + \mathfrak{H} \mid \lambda_1 \mid$. The total spontaneous magnetization I' of the semi-ferromagnetic honeycomb lattice becomes as follows:

$$
\begin{aligned}
I' &= 2m \cdot \lambda_1/\lambda_0 = 2m(\varDelta_1 + \bar{\varDelta}_2) \\
&= 2m\{(\varPsi_0^{+\prime}(s_1 + s_3 + s_5 + \cdots)\varPsi_0^-) + (\varPsi_0^{+\prime} V_1 (s_2 + s_4 + \cdots) V_1^{-1} \varPsi_0^-)\}. \tag{13}
\end{aligned}
$$

Now, in the semi-ferromagnetic honeycomb lattice as shown in Fig. 1, we shall try to interchange the magnetic moment m_0 from ● atoms to ○ atoms. Now the lattice points with ● have no magnetic moment and the lattice points with ○ have the magnetic moment m_0. As before, the total spontaneous magnetization I'' of this semi-ferromagnetic honeycomb lattice is given by

$$
\begin{aligned}
I'' &= 2m \, \lambda_1/\lambda_0 = 2m(\varDelta_2 + \bar{\varDelta}_1) \\
&= 2m\{(\varPsi_0^{+\prime}(s_2 + s_4 + \cdots)\varPsi_0^-) + (\varPsi_0^{+\prime} V_1 (s_1 + s_3 + \cdots) V_1^{-1} \varPsi_0^-)\}. \tag{15}
\end{aligned}
$$

Thus we get from (13) and (15)

$$
\begin{aligned}
I' + I'' &= 2m\{(\varPsi_0^{+\prime}(s_1 + s_2 + s_3 + s_4 + \cdots)\varPsi_0^-) + (\varPsi_0^{+\prime} V_1 (s_1 + s_2 + s_3 + s_4 + \cdots) V_1^{-1} \varPsi_0^-)\}. \\
&= 2m\{\varDelta_1 + \varDelta_2 + \bar{\varDelta}_1 + \bar{\varDelta}_2\}. \tag{16}
\end{aligned}
$$

Since we can arrive at the right hand side of (16) by replacing the operator:

$V_H = e^{\mathfrak{H}(s_1+s_2+s_3+\cdots)}$ instead of $V_{H,1}$, $V_{H,2}$ in (4), $I'+I''$ equals to the total spontaneous magnetization : I of the ordinary ferromagnetic honeycomb lattice in which all atoms have the magnetic moment m_0 and the magnetic energy is given by $\mathfrak{H}(\mu_1+\mu_2+\mu_3+\cdots)+\cdots$ in the presence of the external field.

$$I = I' + I''. \tag{17}$$

Here the semi-ferromagnetic honeycomb lattice which gives the total spontaneous magnetization I' is equivalent completely to the lattice which gives I''. Therefore I' must be equal to I''.

$$I = 2I'. \tag{18}$$

Thus we can obtain the spontaneous magnetization of the ferromgnetic honeycomb lattice from that of the semi-ferromagnetic honeycomb lattice by (18).

§ 3.   The spontaneous magnetization of honeycomb lattice

In order to obtain the spontaneous magnetization of the honeycomb lattice by transforming Potts' result for the triangular lattice, we consider the semi-ferromagnetic honeycomb lattice as shown by Fig. 3. By summing over the spin variables with respect to the vertices of one sublattice which have not the magnetic moment, in the partition function of this lattice (star-triangle transformation), we can get the relation between the partition function of the semi-ferromagnetic honeycomb lattice and that of the ferromagnetic triangular lattice. The interaction parmeters are H for the honeycomb lattice, T for the triangular lattice and the magnetic parameter is $\mathfrak{H}$.
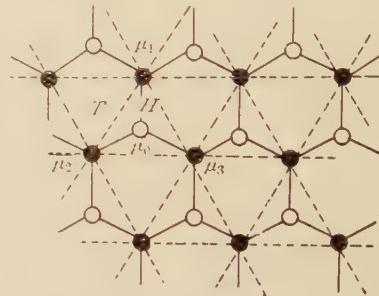
Fig. 3

$$f'_{\text{Honey}}(H,\mathfrak{H}) = \sum_{\{\mu_i=\pm1\}} e^{H\Sigma\mu_0(\mu_1+\mu_2+\mu_3)} \cdot e^{\mathfrak{H}\Sigma(\mu_1+\mu_2+\mu_3)}$$

$$= \sum_{\{\mu_i=\pm1\}} H\, 2\mathrm{ch}\, H(\mu_1+\mu_2+\mu_3) \cdot e^{\mathfrak{H}\Sigma(\mu_1+\mu_2+\mu_3)}$$

$$= A^{\frac{N}{2}} \sum_{\{\mu_i=\pm1\}} e^{T\Sigma(\mu_1\mu_2+\mu_2\mu_3+\mu_3\mu_1)} \cdot e^{\mathfrak{H}\Sigma(\mu_1+\mu_2+\mu_3)}, \tag{19}$$

where

$$A = 2^2 \cdot (\mathrm{ch}\, 3H \cdot \mathrm{ch}^3 H)^{1/2}, \quad e^{4T} = \mathrm{ch}\, 3H/\mathrm{ch}\, H = 2\,\mathrm{ch}\, 2H - 1. \tag{20}$$

Therefore we have

$$f'_{\text{Honey}}(H,\mathfrak{H}) = 2^N (\mathrm{ch}\, 3H \cdot \mathrm{ch}^3 H)^{N/4} \cdot f_{\text{Tri}}(T,\mathfrak{H}) \tag{21}$$

and

$$\frac{1}{N/2} \cdot \frac{\partial}{\partial\mathfrak{H}} \log f'_{\text{Honey}}(H,\mathfrak{H})\bigg|_{\mathfrak{H}\to0} = \frac{1}{N/2} \frac{\partial}{\partial\mathfrak{H}} \log f_{\text{Tri}}(T,\mathfrak{H})\bigg|_{\mathfrak{H}\to0}. \tag{22}$$

The left hand side of (22) gives the spontaneous magnetization per atom by (18) for the honeycomb lattice and the right hand side gives the spontaneous magnetization per atom for the triangular lattice. (The number of the lattice point of the honeycomb lattice is N and that of the triangular lattice is $N/2$.)

$$I_{\text{Honey}}(H) = 2/N \cdot \partial \log f'_{\text{Honey}}(H, \mathfrak{H})/\partial\mathfrak{H}\,|_{\mathfrak{H}\to 0} = I_{\text{Tri}}(T) = I_{\text{Tri}}(T\{H\}). \quad (23)$$

$I_{\text{Tri}}(T)$ is given by Potts[3] as follows.

$$I_{\text{Tri}}(T) = [\,1 - 16y^6/(1+3y^2)(1-y^2)^3\,]^{1/8}, \quad y = e^{-2T}. \quad (24)$$

Thus the spontaneous magnetization per atom for the honeycomb lattice is given by (20), (23) and (24).

$$I_{\text{Honey}}(H) = [\,1 - 16\,z^3(1+z^3)/(1-z)^3(1-z^2)^3\,]^{1/n}, \quad z = e^{-2H}. \quad (25)$$

At low temperature, this formula gives the series solution

$$I_{\text{Honey}}(H) = 1 - 2z^3 - 6z^4 - 18z^5 - 54z^6 - 168z^7 - 534z^8 \cdots \quad (26)$$

This result coincides with the rigorous series solution at low temperature[4].

§ 4. The spontaneous magnetization of Kagomé lattice

In order to obtain the spontaneous magnetization of the kagomé lattice by transforming the result for the honeycomb lattice, we consider the so-called decorated honeycomb lattice as shown by Fig. 4. This too is a kind of semi-ferromagnetic lattice since the decorated ● atoms have the magnetic moment m_0 and the ○ atoms of the vertices have no magnetic moment. The interaction parameters are H for the honeycomb, K for the kagomé and l for the decorated honeycomb lattice. The magnetic parameters are $\mathfrak{H}$ for the kagomé, decorated honeycomb and $\mathfrak{H}^*$ for the honeycomb lattice. By summing at first over the spin variables with respect to the vertices in Fig. 4, we can find the relation between the partition function of the decorated honeycomb lattice and that of the kagomé lattice (star-triangle transformation).

Fig. 4

$$f_{\text{Dec}}(l, \mathfrak{H}) = \sum_{[\mu]=\pm 1} e^{l\mu_0(\mu_1+\mu_2+\mu_3)+\cdots} \cdot e^{\mathfrak{H}(\mu_1+\mu_2+\mu_3)+\cdots}$$

$$= \sum_{[\mu]'=\pm 1} \Pi\, 2\text{ch}\, l(\mu_1+\mu_2+\mu_3) \cdot e^{\mathfrak{H}(\mu_1+\mu_2+\mu_3)+\cdots}$$
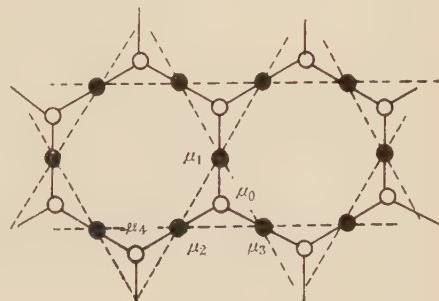
$$= A^N \sum_{[\mu]'=\pm 1} e^{K(\mu_1\mu_2+\mu_2\mu_3+\mu_3\mu_1)+\cdots} \cdot e^{\mathfrak{H}(\mu_1+\mu_2+\mu_3)+\cdots}$$

$$= A^N \cdot f_{\text{kagomé}}(K, \mathfrak{H}). \quad (27)$$

where

$$A = 2^2 \cdot (\text{ch } 3l \cdot \text{ch}^3 l)^{1/2}, \quad e^{4K} = \text{ch } 3l / \text{ch } l = 2 \text{ ch } 2l - 1. \tag{28}$$

Also, by summing at first over the spin variables with respect to the decorated atoms in Fig. 4, we can find the relation between the partition function of the decorated honeycomb lattice and that of the honeycomb lattice. (iteration process)

$$
\begin{aligned}
f_{\text{Dec}}(l, \mathfrak{H}) &= \sum_{[\mu] = \pm 1} e^{l\mu_2(\mu_0 + \mu_4) + \cdots} \cdot e^{\mathfrak{H}(\mu_1 + \mu_2 + \mu_3) + \cdots} \\
&= \sum_{[\mu] = \pm 1} e^{l\mu_2(\mu_0 + \mu_4) + \mathfrak{H}\mu_2 + \cdots} \\
&= \sum_{[\mu'']= \pm 1} \prod 2 \text{ ch } \{l(\mu_0 + \mu_4) + \mathfrak{H}\} \\
&= \sum_{[\mu'']= \pm 1} B^{3/2N} \cdot e^{H\mu_0\mu_4 + \cdots} \cdot e^{3\mathfrak{H}*(\mu_0 + \mu_4) + \cdots} \\
&= B^{3/2N} \cdot f_{\text{Honey}}(H, 3\mathfrak{H}*), \tag{29}
\end{aligned}
$$

where

$$B = 2 \{\text{ch}(2l + \mathfrak{H}) \text{ch}(2l - \mathfrak{H}) \text{ch}^2 \mathfrak{H}\}^{1/4},$$

$$4H = \log \text{ch}(2l + \mathfrak{H}) + \log \text{ch}(2l - \mathfrak{H}) - 2 \log \text{ch } \mathfrak{H},$$

$$4\mathfrak{H}* = \log \text{ch}(2l + \mathfrak{H}) - \log \text{ch}(2l - \mathfrak{H}). \tag{30}$$

Thus we can obtain the relation between the partition function of the honeycomb lattice and that of the kagomé lattice from (27) and (29),

$$f_{\text{kagomé}}(K, \mathfrak{H}) = \frac{\{\text{ch}(2l + \mathfrak{H}) \text{ch}(2l - \mathfrak{H}) \text{ch}^2 \mathfrak{H}\}^{3/8N}}{2^{N/2} \cdot [\text{ch } 3l \cdot \text{ch}^3 l]^{N/2}} \cdot f_{\text{Honey}}(H, 3\mathfrak{H}*). \tag{31}$$

The next relation are obtained from (30).

$$4 \, \partial H / \partial \mathfrak{H} = \text{th}(2l + \mathfrak{H}) - \text{th}(2l - \mathfrak{H}) - 2\text{th } \mathfrak{H}, \quad \partial H / \partial \mathfrak{H} |_{\mathfrak{H} \to 0} = 0,$$

$$4 \, \partial \mathfrak{H}* / \partial \mathfrak{H} = \text{th}(2l + \mathfrak{H}) + \text{th}(2l - \mathfrak{H}), \quad \partial \mathfrak{H}* / \partial \mathfrak{H} |_{\mathfrak{H} \to 0} = 1/2 \cdot \text{th } 2l, \tag{32}$$

$$\mathfrak{H}* \to 0 \quad \text{with} \quad \mathfrak{H} \to 0.$$

From (31) and (32), one sees that the spontaneous magnetization of the kagomé lattice and that of the honeycomb lattice are related with each other in the following manner.

$$I_{\text{kagomé}}(K) = \text{th } 2l \cdot I_{\text{Honey}}(H) = \text{th}(2l\{K\}) \cdot I_{\text{Honey}}(H\{l[K]\}). \tag{33}$$

Thus we can obtain the spontaneous magnetization of the kagomé lattice per atom from (25), (28), (30) and (33) as follows:

$$I_{\text{kagomé}}(K) = \frac{(1 + 3u^2)^{1/2}(1 - u^2)^{1/2}}{(1 + u^2)} \left[1 - \frac{128u^6(1 + u^2)^3(1 + 3u^4)}{(1 - u^2)^6(1 + 3u^2)^2} \right]^{1/8}. \tag{34}$$

$$u = e^{-2K}.$$

At low temperature, this formula gives the series solution

$$I_{\text{Kagomé}}(K) = 1 - 2u^4 - 12u^6 - 56u^8 - 240u^{10} - 1353u^{12} \cdots. \tag{35}$$

This result coincides with the rigorous series solution at low temperature.[4]

5. Comparison of the results

The results so far obtained are shown in the table I. As shown in the table I, the square lattice and the kagomé lattice with the same coordination number $Z=4$ are different in the third and higher terms in the low temperature series expansion. Near the Curie temperature, I's have the form $[const \cdot (1-x/x_c)]^{1/8}$ and fall down rapidly. These curves are plotted in Fig. 5. ($I-x$ curves) and in Fig. 6 ($I-x/x_c$ curves).

Table I

Lattice type	Interaction parameter	Curie point	Spontaneous magnetization	Near the Curie temperature
Square	$x = e^{-2L}$	$x_c = \sqrt{2} - 1$ $\fallingdotseq 0.4142$	$I_S = \left[1 - \dfrac{16x^4}{(1-x^2)^4}\right]^{1/8}$: Yang	$I_S \sim [4\sqrt{2}\,(1-x/x_c)]^{1/8}$ $\fallingdotseq [5.657(1-x/x_c)]^{1/8}$
Triangular	$y = e^{-2T}$	$y_c = 1/\sqrt{3}$ $\fallingdotseq 0.5773$	$I_T = \left[1 - \dfrac{16y^6}{(1+3y^2)(1-y^2)^3}\right]^{1/8}$: Potts	$I_T \sim [8(1-y/y_c)]^{1/8}$
Honeycomb	$z = e^{-2H}$	$z_c = 2 - \sqrt{3}$ $\fallingdotseq 0.2680$	$I_H = \left[1 - \dfrac{16z^3(1+z^3)}{(1-z)^3(1-z^2)^3}\right]^{1/8}$	$I_H \sim [1/14(-15+46\sqrt{3})$ $(1-z/z_c)]^{1/8}$ $\fallingdotseq [4.619(1-z/z_c)]^{1/8}$
Kagomé	$u = e^{-2K}$	$u_c =$ $(2/\sqrt{3}-1)^{1/2}$ $\fallingdotseq 0.3933$	$I_K = \dfrac{(1+3u^2)^{1/2}(1-u^2)^{1/2}}{(1+u^2)}$ $\left[1 - \dfrac{128u^6(1+u^2)^3(1+3u^4)}{(1-u^2)^6(1+3u^2)^2}\right]^{1/8}$	$I_K \sim [144(1403-810\sqrt{3})$ $(1-u/u_c)]^{1/8}$ $\fallingdotseq [5.602(1-u/u_c)]^{1/8}$

Lattice type	Series expansion at low temperature
Square	$I_S = 1 - 2x^4 - 8x^6 - 34x^8 - 152x^{10} - 714x^{12} \cdots$
Triangular	$I_T = 1 - 2y^6 - 12y^{10} + 2y^{12} - 78y^{14} + 24y^{16} - 548y^{18} \cdots$
Honeycomb	$I_H = 1 - 2z^3 - 6z^4 - 18z^5 - 54z^6 - 168z^7 - 534z^8 \cdots$
Kagomé	$I_K = 1 - 2u^4 - 12u^6 - 56u^8 - 240u^{10} - 1353u^{12} \cdots$

As shown in Fig. 6, the curve for the triangular lattice runs the most outside and that for the honeycomb lattice runs the most innerside. Near the absolute temperature zero,

$$I_S = 1 - 2x_c^4(x/x_c)^4 \cdots\cdots \text{ for the square lattice,}$$

$$I_K = 1 - 2u_c^4(u/u_c)^4 \cdots\cdots \text{ for the kagomé lattice,}$$

hence $I_S < I_K$ because $x_c^4 > u_c^4$. Near the Curie temperature, on the other hand,

Fig. 5

Spontaneous magnetization :

$(I-x$ curves)

a : Honeycomb
b : Kagomé
c : Square
d : Triangular
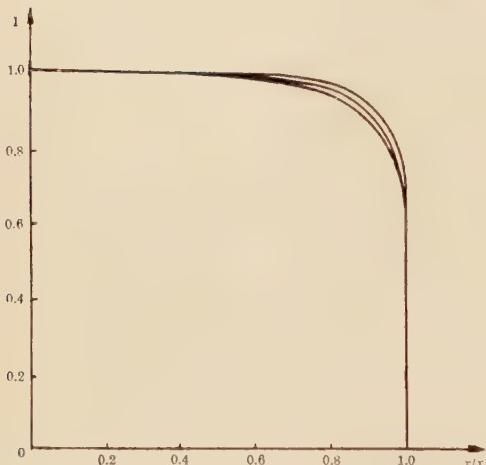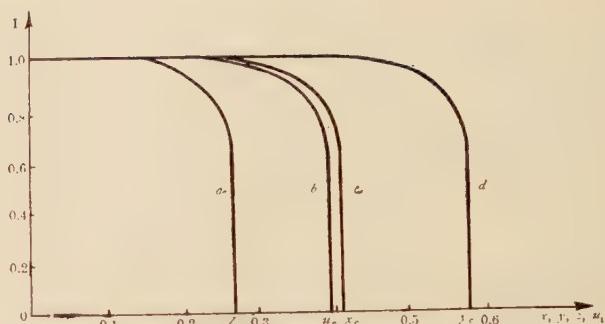
Fig. 6

Spontaneous magnetization :

$(I-x/x_c$ Curves)

From inside, the honeycomb, the square and the triangular latties

$$I_S \simeq [5.657(1-x/x_c)]^{1/8} > I_K \simeq [5.602(1-u/u_c)]^{1/8}.$$

Therefore the curves of the spontaneous magnetizations of the square and kagomé lattices cross at an intermediate temperature. In any case the difference is very small to be represented in the figure.

§ 6. On Potts' relations

In order to generalize Yang's result for the square lattice, Potts pointed out that the spontaneous magnetization of the square lattice can be written in a form different from that given by Yang as follows :

$$I_S{}^8 = 1 - \mathrm{sh}^2\, 2L^+/\mathrm{sh}^2\, 2L, \qquad\qquad (36)$$

where L^+ is obtained from L by the inversion transformation :

$$\mathrm{sh}\, 2L \cdot \mathrm{sh}\, 2L^+ = 1 \qquad (37)$$

Fig. 7

From (36), (37), $I_S{}^8$ becomes as follows :
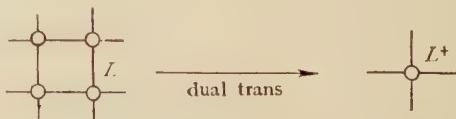
$$I_s^s=1-\text{sh}^2\,2\,L^+/\text{sh}^2 2\,L=1-1/\text{sh}^4 2L=1-16x^4/(1-x^2)^4,\ \ x=e^{-2L}. \qquad (38)$$

This result coincides with Yang's. Potts supposed that the spontaneous magnetization for the triangular lattice may also be obtained in a form analogous to (36):

$$I_T^s=1-\text{sh}^2 2\,T^+/\text{sh}^2 2\,T. \qquad (39)$$

In the isotropic triangular lattice, the inversion transformation $T \to T^{\,5)}$ is given by (see Fig. 8)

$$(e^{4T}-1)(e^{4T+}-1)=4. \qquad (40)$$
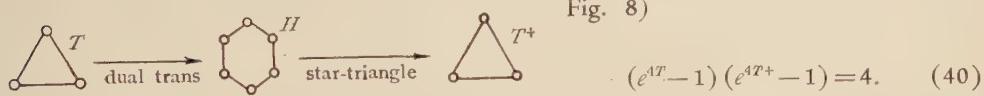
Fig. 8

From (39) and (40), I_T is obtained.

$$I^s_T=1-16y^6/(1-y^2)(1+3y^2),\ \ y=e^{-2T}. \qquad (41)$$

Potts has checked this result by the rigorous solution at low temperature.

We should like to examine whether the Potts' assumption holds always, in other words, whether the spontaneous magnetization for any two dimensional Ising lattices can be written in the form of (36). In the honeycomb lattice, the inversion transformation $H \to H^+$ runs as follows.[5] (see Fig. 9)

$$(\text{sh}2H-1)(\text{sh}2H^+-1=1).$$

$$(42)$$
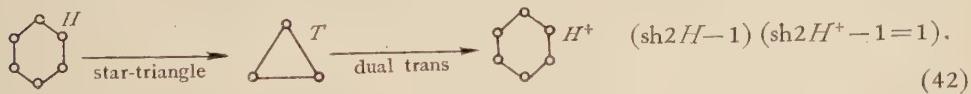
Fig. 9

If Potts' assumption is valid,

$$I_H^s=1-\text{sh}^2 2H^+/\text{sh}^2 2H, \qquad (43)$$

the spontaneous magnetization of the honeycomb lattice becomes as follows :

$$I_H^s=1-16\,z^3\cdot(1+z^3)/(1-z)^3\cdot(1-z^2)^3,\ \ z=e^{-2H}. \qquad (44)$$

This result is in accord with (25) that has been checked by the rigorous series solution at low temperature. In the Kagomé lattice, the inversion transformation is obtained from decorated honeycomb lattice or diced lattice (dual net of kagomé). (see Figs. 10, 11)

$$e^{4K}=2e^{2H}-1$$
$$e^{4K+}=2e^{2H+}-1 \qquad (45)$$
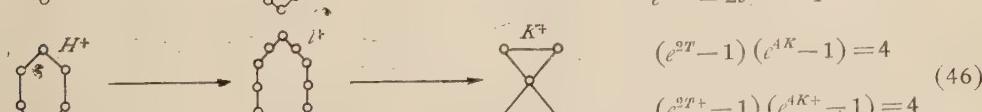
$$(e^{2T}-1)(e^{4K}-1)=4$$
$$(e^{2T+}-1)(e^{4K+}-1)=4 \qquad (46)$$

Fig. 10

Therefore the inversion transformation $K \to K^+$ is obtained from (45) or (46) because $H \to H^+$ or $T \to T^+$ are obtained from (42) or (40). (of course, (45) and (46) give the

Fig 11

same result)

$$(c^{4K}-1)^2(c^{4K+}-1)^2/(c^{4K}+1)(c^{4K+}+1)$$
$$=16 . \tag{47}$$
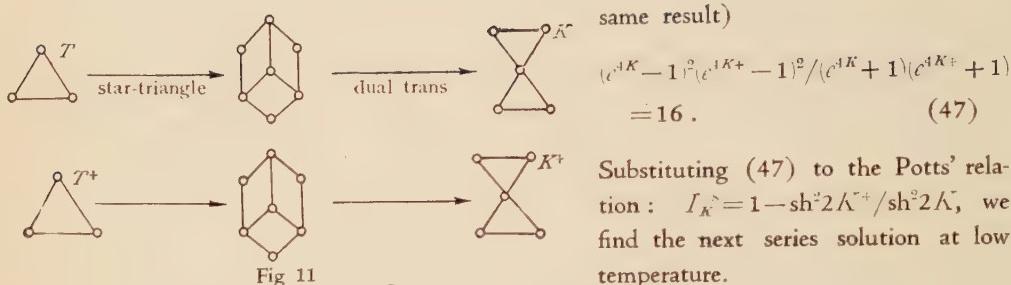
Substituting (47) to the Potts' relation : $I_{K'}=1-\mathrm{sh}^2 2K^+/\mathrm{sh}^2 2K'$, we find the next series solution at low temperature.

$$I_K{}^8=1-\mathrm{sh}^2 2K^+/\mathrm{sh}^2 2K=1-32u^4-64\sqrt{2}\,u^5-544u^6$$
$$-1504\sqrt{2}\,u^7-9664\,u^8\cdots\cdots . \tag{48}$$

On the other hand, $I_K{}^8$ obtained from (34) is

$$I_K{}^8=1-16\,u^4-96\,u^6-336\,u^8\cdots\cdots . \tag{49}$$

The series solution (48) and (49) do not give the same result. Of course, (34) is what has been checked already by the rigorous series expansion at low temperature. Therefore the formula of the spontaneous magnetization gneralized by Potts is suitable for the square, triangular and honeycomb lattice, but not for the kagomé lattice. This situation is like to the Onsager's relation for the Curie temperature :

$$\mathrm{ch}\,2I_c=\sec(\pi/z) \tag{50}$$

that has been adequate only for the square, triangular and honeycomb lattice, but not for the kagomé lattice.[6]

In conclusion, we should like to express our sincerest thanks to Prof. Kôdi Husimi and Mr. Itiro Syôzi for their kind interest in this work.

References

1) L. Onsager, Phys. Rev. **65** (1944), 117 ; B. Kaufman, Phys. Rev. **76** (1949), 1232 ; Y. Nambu, Prog. Theor. Phys. **5** (1950), 1 ; K. Husimi and I. Syozi, Prog. Theor. Phys. **5** (1950), 117, 341 ; H. N. V. Temperley, Proc. Roy. Soc. A **202** (1950), 202 ; R. M. F. Houtappel, Physica **16** (1950), 425 ; G. H. Wannier, Phys. Rev. **79** (1950), 357 ; G. F. Newell, Phys. Rev. **79** (1950), 877.

2) C. N. Yang, Phys. Rev. **85** (1952), 808.

3) R. B. Potts, Phys. Rev. **88** (1952), 392.

4) Although the rigorous series expansion at low temperature for the kagomé lattice has not yet been obtained, we have checked directly by performing the rigorous series expansion for the honeycomb and kagomé lattice by the method of Oguchi. (Jour. Phys. Soc. Japan **6** (1951)).

5) for example, Quantum statistical mechanics, by K. Husimi (in Japanese).

6) I. Syozi, Busseiron-kenkyu **40** (1951) (in Japanese) ; K. Kano and S. Naya, Busseiron-kenkyu **40** (1951) (in Japanese).

Nucleon Scattering by a Potential

Masahiko MATSUMOTO* and Wataro WATARI

*Department of Industrial Chemistry, Faculty of Engineering,
Kyoto University*

(Received November 2, 1953)

Scattering of a nucleon by a central potential is discussed. The potential consists of a sequence of square wells. Relations between the phase shift of the partial wave and the potential shape are extensively examined for the incident energy lower than 300 Mev in laboratory system. The S, P, D, F and G partial waves are taken into account. The results are summarized in the form of graphs.

Introduction

Several authors made theoretical discussions on nuclear forces since the publication of the experimental results of the neutron-proton and the proton-proton scattering at high energy. Many of these investigations,[1,2,3,4] however, are phenomenological in the sense that they tried to explain the experimental results on the basis of assumed potential models. On the other hand, the nuclear force must be derived from the π-meson theory, although the theory of pion-nucleon interaction at present has evidently several defect to be improved in future. At the present stage of the theory, one must expect some new knowledge to be added to his present one about nucleons and their interaction.

In these circumstances, Taketani, Nakamura and Sasaki[5] proposed the following method as promising. The static π-meson potential is assumed in the outside region, while in the inside region, the phenomenological potential, represented by the square well, is adopted. The square well is so adjusted as to fit each experimental result, and might be varied, if necessary, according to each process and energy region concerned. So far the low energy problem was treated, following this method, by Taketani, Machida and Ohnuma[6] and Brueckner and Watson[7] using the pseudoscalar meson theory. But high energy problems have not yet been treated.

Reflecting on these facts, we think that it is still necessary to have a look over following two points :

1) *The relation between the collision energy and the extent to which each partial wave are affected by some part of potential.*

If, at given energy, some partial waves with higher angular momentum are scarcely affected by the inside region of the potential, then the phase shifts of these waves are

---

* On leave from Shiga University.

substantially determined from the meson theoretical potential when one adopts Taketani's prescription. And we may check the applicability of the meson theory by comparing theoretical results given by the meson theoretical outside potential with the phenomena to which only phase shifts of some higher waves contribute. As regards to the qualitative aspects of this problem, it is commonly known that waves with higher angular momentum are not affected by a short range potential at low energy. But in order to push forward the Taketani method, one must make such an argument as above quantitatively more precise.

2) *The relation between the phase shifts of the scattered waves and the typical characteristics of the potential.*

Here the typical characteristics of the potential means, for example, the depth and the width of the tail of the potential or whether it has a hard core about its origin and so on. Of course, it is well known that the potentials with long tail[1] or hard core[1] is necessary for the best fit to the observed angular distribution. It is yet necessary to study quantitatively the relation mentioned above, because the main features of the potential are determined by Taketani's method but the detailed features must be determined phenomenologically to reproduce the experimental data. So one needs to look over this point to prepare for future developments.

In the present paper, the scatterieg due to only the central potentials is considered as the first step of the investigation along this line. S, P, D, F and G partial waves (i.e., waves with angular momentum $l=0$, 1, 2, 3, 4), and the energy below 300 Mev in laboratory system are taken into account. In section 1, the change of the phase shift for each partial wave will be investigated, which is caused by the variation of some part of potential, e.g. the radius of the hard core or the depth of the outer part. But no attempt is made there to explain the present experimental results by means of these trial potentials. In section 2, using the potentials adjusted to the low energy scattering experiments, the phase shifts of each partial wave are plotted as the functions of the collision energy.

The generalized " L-shape potential "[*] is exclusively used. This potential consists of a sequence of square wells with different depth and range. The reasons for adopting this shape of potential are the followings. Firstly, in this case of the potential, the Schroedinger equation can be solved exactly and the solution is represented by the spherical Bessel functions, therefore, the trouble of calculations can be extremely cut down. All phase shifts are obtained by the exact calculation. Secondly, the potential shape may not be so delicate as to give serious effects to the results such as cross sections or phase shifts in the scattering problem. So we feel that this " L-shape " potential is sufficiently useful for our purpose. This point is examined by comparing the L-shape potential with the exponential or the Yukawa potential.

The numerical results are summarized in the form of graphs. The extensive use of these graphs will serve the purpose of the estimation of phase shifts caused by the potential of arbitrary shape. Effects of the inner part of potential are especially examined. It will be shown that the waves of the impact parameter a are not affected by any change of potentials inner than $a/2$, and they are mainly scattered by the outer region of potentials
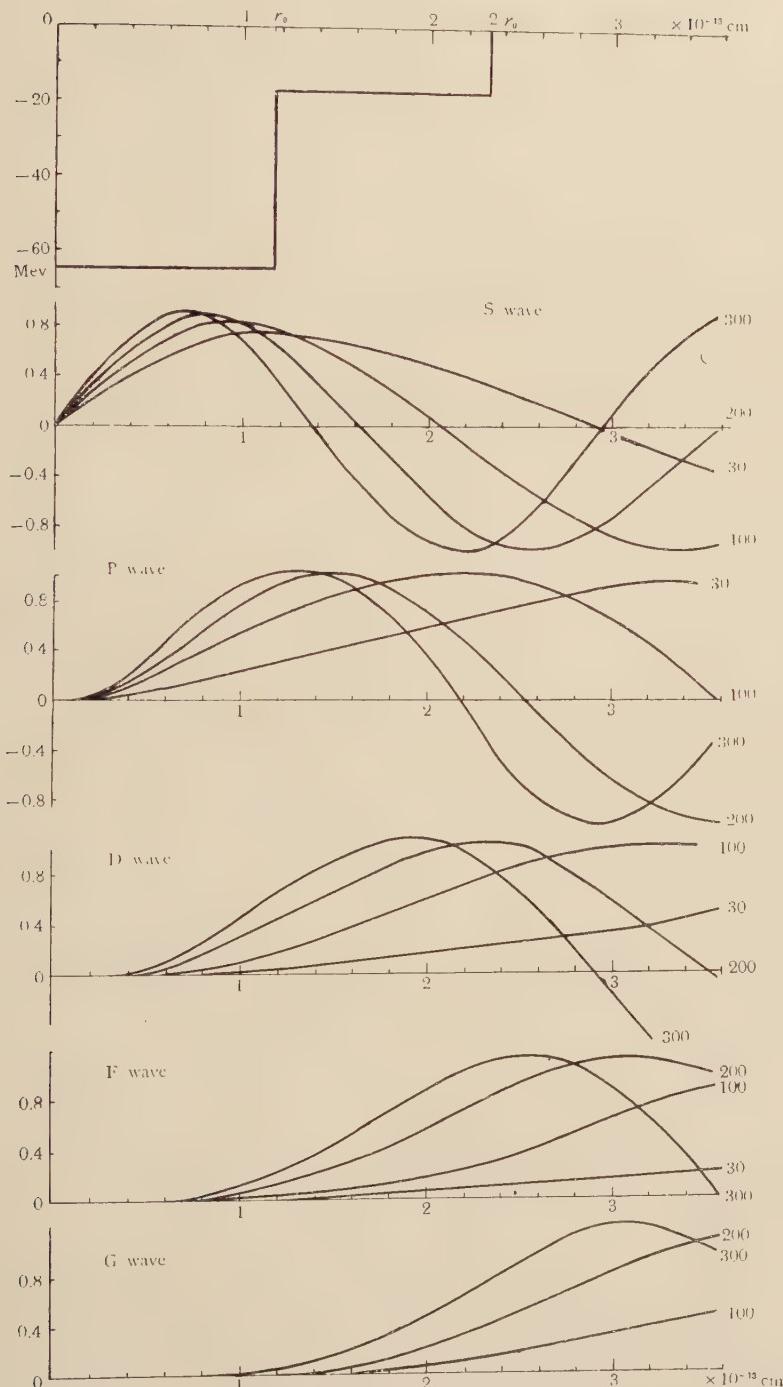
Fig. 1. The standard potential and partial waves

For potentials, Abscissa : length in units 10^{-13}cm. ; Ordinate : Depth of potential. (in Mev)
For wave functions, Abscissa is the same with that for the potential. Each wave is normalized
so that its amplitude becomes 1 at infinity.

than a.   For example, if the collision energy is lower than 50 Mev, the inside phenome-
nological potential is not affected on the P and higher partial waves, and if the inner part
of the potential for P wave does not so strongly attractive (not deeper than -200 Mev),
P wave is not affected by the inside region of the potential when the collision energy is
lower than 100 Mev.   In the actual case, there is no bound ^3P state, so that above
condition may be satisfied.   Then, we may say that, if the collision energy is lower than
100 Mev, only S-wave has phenomenological character, and the phase shifts of P and
higher partial waves are substantially determined from the outside meson potentials.

The following units are used throughout this paper : Mev for energy, 10^{-13} cm. for
length and radian for phase shifts.

§ 1.   Relations between potential shape and phase shifts
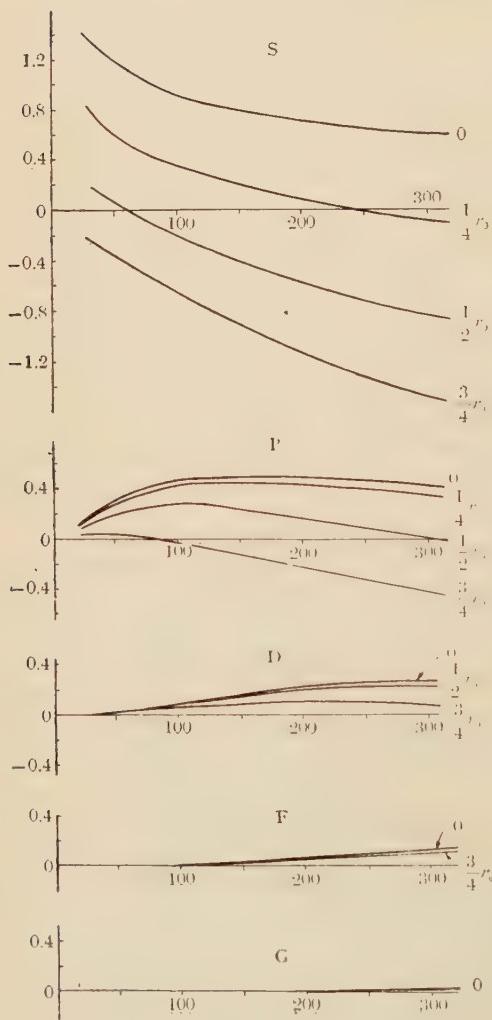
It is supposed that the incident beam
of the impact parameter a can approach
up to a from the centre of the scatterer.[9]
Therefore, the incident wave of azimuthal
quantum number l and de Broglie wave
number k is mainly scattered by the outer
part of the potential than $a=l/k$.   In the
case of usual potentials, this intuitive
argument holds with quantitative accuracy as
can be seen from the following results.
Value of a for each partial wave is shown
in Table I.   The behavior of phase shift
in the low energy region differs according
to whether potential gives bound states or
not.   This has already been discussed by
Mott and Massey[10], and this point is not
discussed any more and only the cases
where the collision energy is higher than
about 30 Mev is examined.

Fig. 2.   The phase shift of each partial wave is
plotted as the function of the incident energy with
several different hard core radius.   Abscissa : incident
energy in laboratory system in Mev ; Ordinate : phase
shift in radian.   The number attached to each curve
is the radius of hard core, where $r_0 = 1.16 \times 10^{-13}$ cm,

Table 1   Table of the impact parameters of various partial waves.

Energy in *L.S.* (Mev)	30	50	100	200	300
k (10^{13}cm^{-1}) l	0.602	0.777	1.064	1.553	1.903
1 *P*-wave	1.661	1.287	0.940	0.644	0.525
2 *D*-wave	3.322	2.574	1.880	1.288	1.052
3 *F*-wave	4.983	3.861	2.820	1.932	1.578
4 *G*-wave	6.644	5.148	3.760	2.576	2.104
5 *H*-wave				3.220	2.630
6				3.864	3.156
7					3.682

The potential shown in Fig. 1 is adopted as the standard one. This standard potential is adjusted to the low energy experimental value of triplet data (the effective range and the scattering length) and gives the similar phase shift as the one derived from the exponential well, in a wide region of energy. The shape of each partial wave which is scattered by this potential is also shown in Fig. 1 when the incident energy is 30 Mev, 100 Mev, 200 Mev, and 300 Mev in laboratory system.
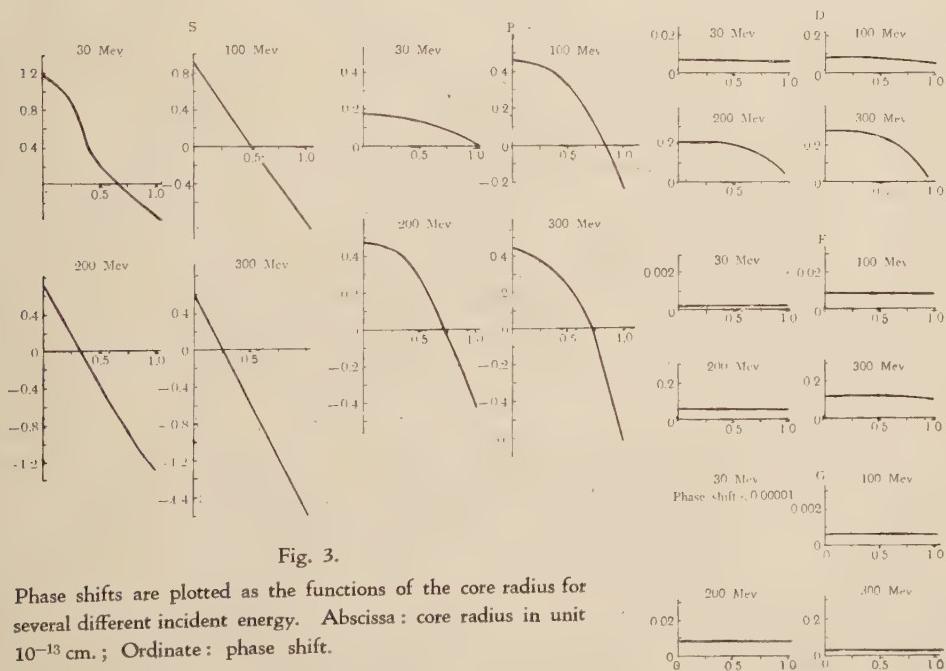
Fig. 3.

Phase shifts are plotted as the functions of the core radius for several different incident energy. Abscissa : core radius in unit 10^{-13} cm.; Ordinate : phase shift.

(i)   *Variation of inner part of the potential well*

Jastrow[4] has proposed to replace the central part of nuclear potential inner than the nucleon Compton wave length λ or its twice 2λ by a so-called hard core. This idea as

M. Matsumoto and W. Watari
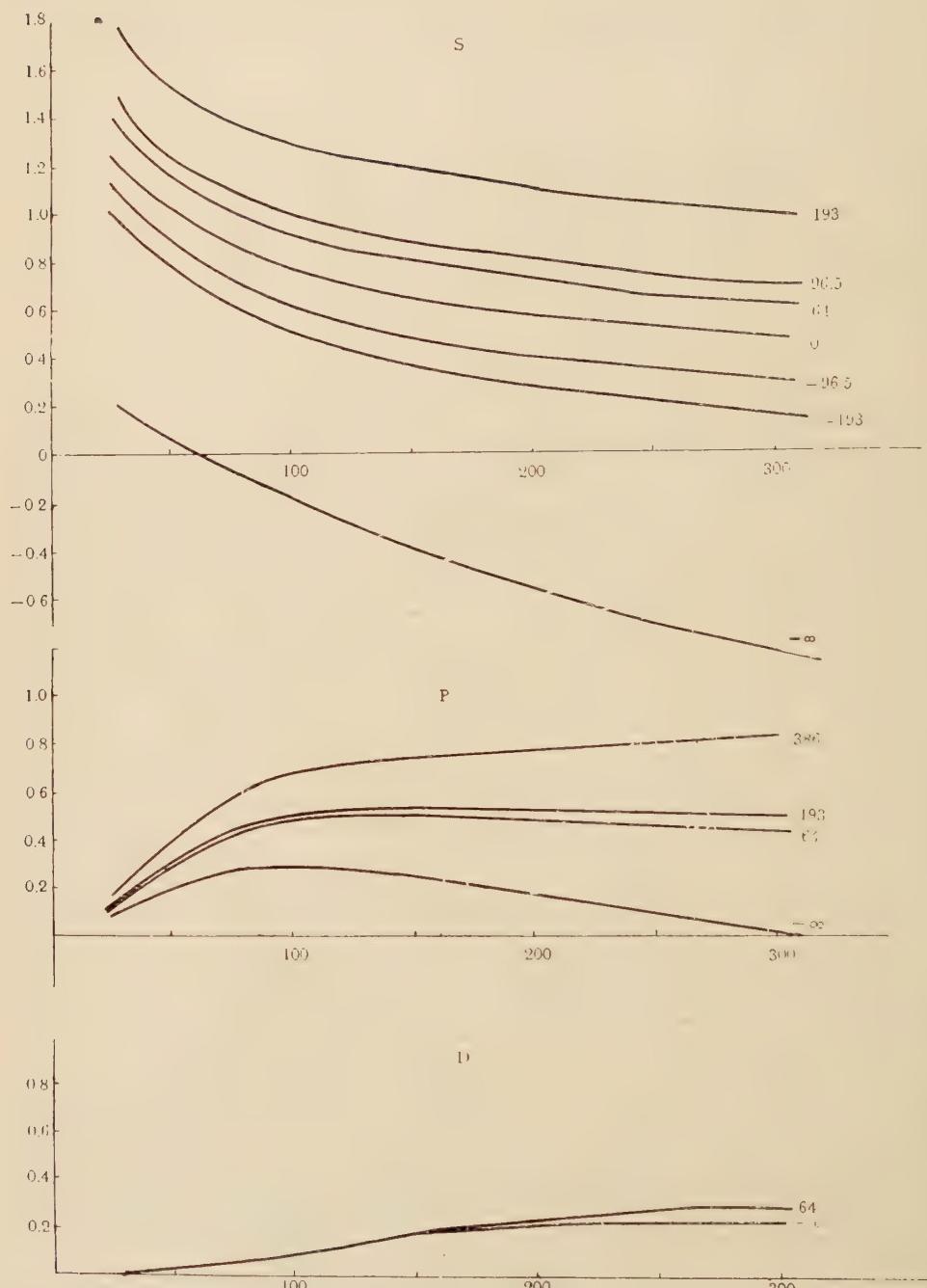
Fig. 4. The phase shift of each partial wave is plotted as the function of the incident energy with several different depth of inner part of potential. Abscissa : incident energy ; Ordinate : phase shift. The number attached to each curve is the depth of the inner part in unit Mev.

adopted in the potential derived from meson theory can represent many complicated effects such as the higher order effects, the non-adiabatic effects, and so on. In this sub-section it is investigated what influences due to some change of the inner potential ($r \ll$ meson Compton wave length) are induced on phase shifts.

(a)   In Figs. 2 and 3, the change of phase shifts due to the variation of the radius of *hard core* are shown. The potentials are constructed by adding the hard core to the standard potential. The number attached to each curve in Fig. 2 shows the radius of hard core where $r_0 = 1.16 \times 10^{-13}$cm. This graph does not represent only pure hard core effect but it contains also the effect caused by the reduction of the attractive potential which is cut off by the hard core. That is, the result of this sub-section depends on the shape of the attractive potential. From the results given here, it seems that each partial wave will hardly receive any influence of hard core, if the hard core radius is shorter than half of a for that wave. This is the effect of the centrifugal force concealing the hard core effect. In other words the amplitude of the wave near $r=0$ is not large enough for the wave to be affected by the hard core. Therefore, we may conclude that effects of any change of the potential are negligible when it takes place in this inner region than $a/2$.* In the case of the large incident energy, the S-wave phase shift is proportional to the hard core radius times the wave number k.

(b)   The behavior of the phase shift due to the change of the *potential depth* inner than $r_0/2$ of the standard potential is investigated, and graphs similar to the case of (a) are obtained. Phase shifts are plotted against energy and potential depth in Fig. 4 and Fig. 5 respectively. The number attached to each curve of figures in Fig. 4 is the value of the depth of potential centre in Mev units. F and G waves receive no influence of this potential part as it will be seen from Figs. 2 and 3. D-wave also receives only little influence.
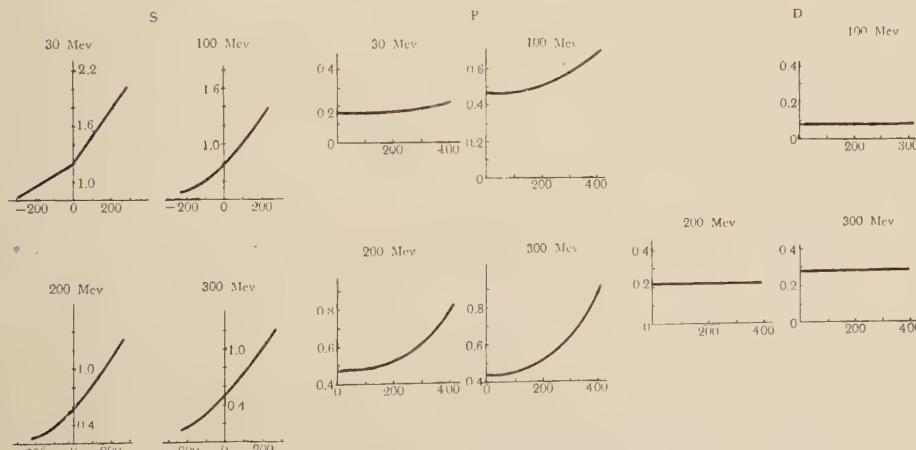
Fig. 5.   Phass shifts are plotted as the functions of the depth of the inner part of the potential. Abscissa : depth of potential in unit Mev; Ordinate : phase shift.

---

*)   Of course, there occurs the exceptional case when a very strong *attractive* potential is added to this region.

From results of (a) and (b), the effect resulted from any change of the inner part of some potential will quantitatively be estimated to some extent.
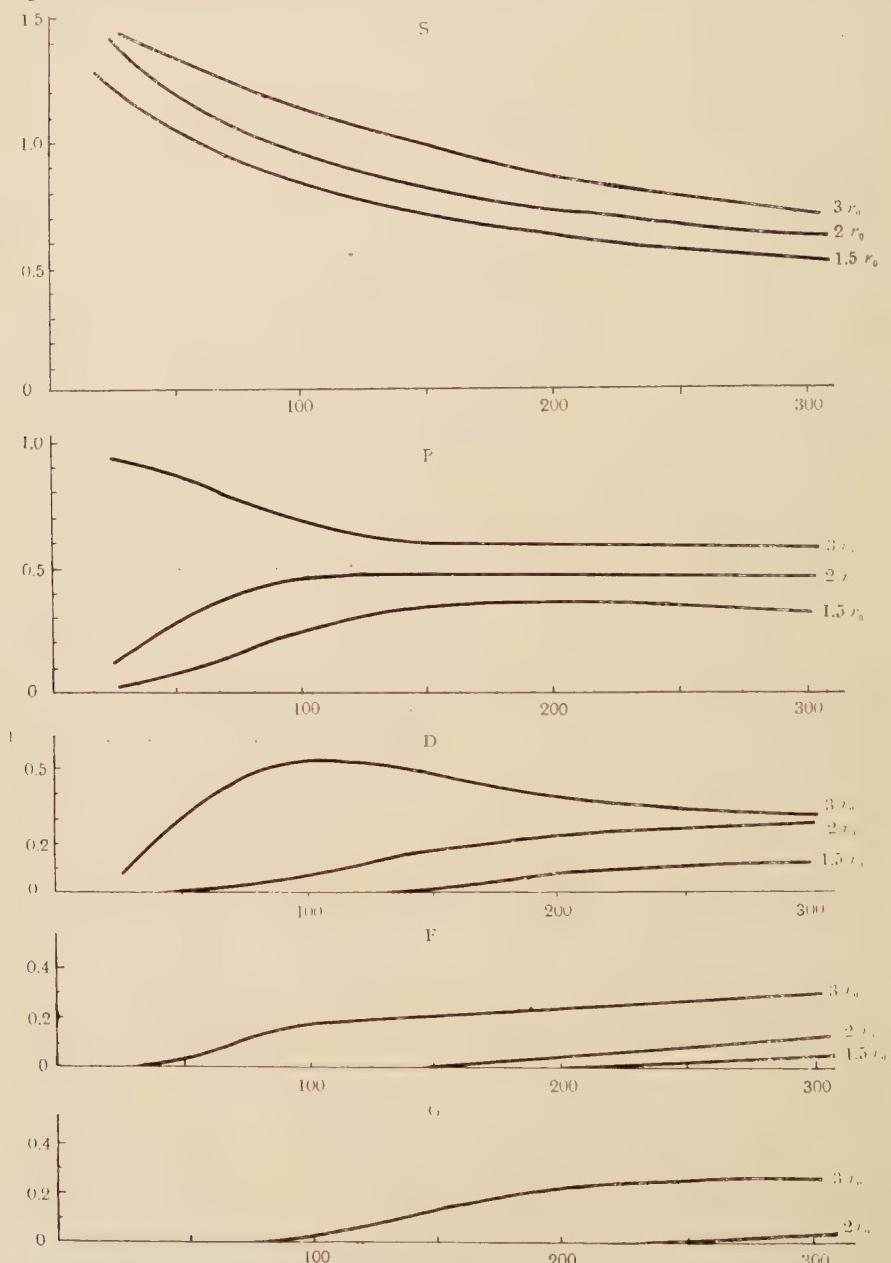
Fig. 6.

The behavior of phase shifts are shown as the function of the incident energy with several different length of potential tail. The number attached to each curve is the coordinate of the end of tail. Abscissa: incident energy; Ordinate: phase shift.
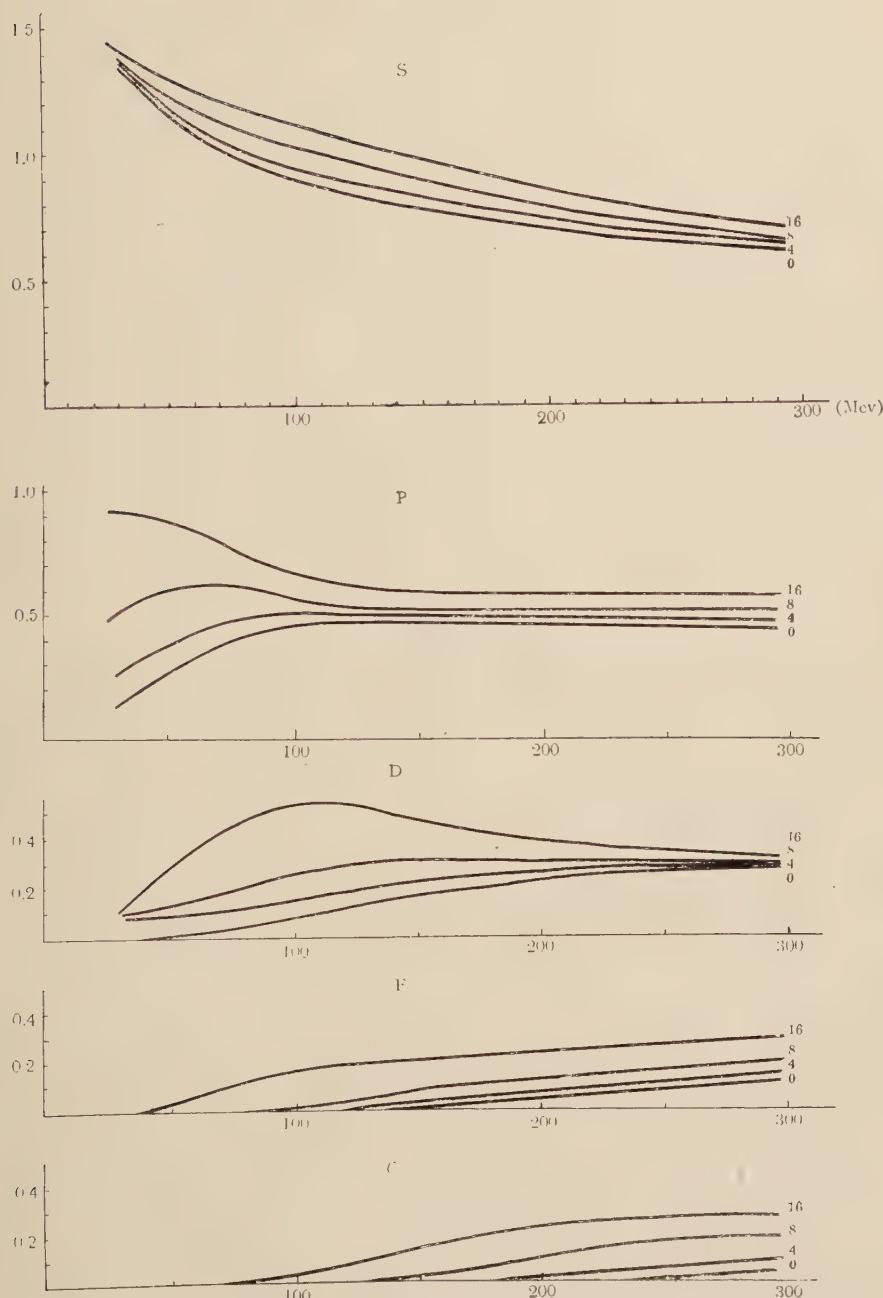
Fig. 7.

Phase shifts are plotted as the functions of the incident energy with some different depth of the tail of potential.   Abscissa : incident energy in Mev. ; Ordinate : phase shift in radian.

(ii) *Variation of outer part of the potential well*

The following graphs are presented here in order to see the change of phase shifts as the outer part of the potential $(r > r_0)$ is varied.

(a) Firstly, the behavior of the phase shift is examined in cases where the potential tail with fixed depth stretches to a *various range*. (Fig. 6). A number attached to each curve represents the range of potential " tail ". That is, $2r_0$ corresponds to the standard potential, $1.5r_0$ and $3r_0$ are the potential obtained by cutting or stretching the tail of standard potential at $r = 1.5r_0$ and to $r = 3r_0$ respectively.

(b) The standard potential is connected here with a tail stretches from $2r_0$ to $3r_0$, and the *depth* of the tail is varied from 0 to 16 Mev. The behavior of phase shift with this change of the potential is illustrated in Fig. 7.

According to above illustration, the influence due to the change of the inner part of potential chiefly appears as the change of S-wave phase shift. On the contrary, the change of the outer part of potential affects the large change to the phase shift of high angular momentum wave, especially in low incident energy regions. Thus we may summarize our results in the statement that, briefly speaking, the inner part of the potential contribute to total cross section while the angular distribution depends on the outer part of the potential.

§ 2. The case with the potentials adjusted for the low energy experiments

In the former section, various natures of potential was studied by varying the potential shape without regards to their fit to the experimental results. Although they are independent of the practical problems, the way followed there is suitable for the purpose of clarifying the situation underlying our analysis. Then, in this section, let us review again the whole cases of the potential's shape, thereby restricting them which are adjusted to low energy experiments. Every potentials here adopted are adjusted so as to give correctly[11] :
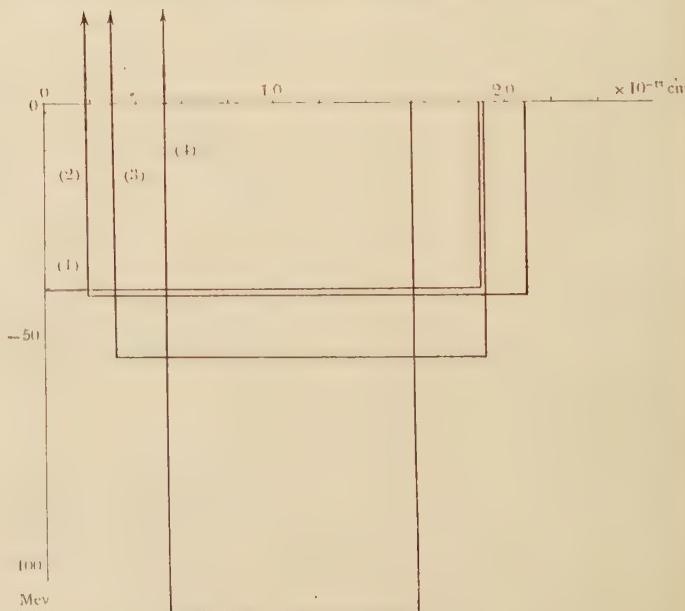
Fig. 8. The triplet potential with hard core.
Abscissa : length ; Ordinate : depth of potential.

The triplet effective range

$$r_{ot} = (1.70 \pm 0.03) \times 10^{-13} \text{cm},$$

the triplet scattering length

or
$$a_t = (5.378 \pm 0.023) \times 10^{-13} \text{cm.},$$

the singlet effective range

$$r_{os} = (2 \cdot 7 \pm 0.5) \times 10^{-13} \text{cm.},$$

the singlet scattering length

$$a_s = -(-23.69 \pm 0.06) \times 10^{-13} \text{cm.}$$

(i) *Triplet potentials*

(a) Potentials shown in Fig. 8 are all adjusted to give the triplet parameters but have different radius of *hard core*. The larger the hard core radius becomes the deeper and more narrow becomes the attractive potential well. This change is necessary to keep the phase shift of triplet *S*-wave to fit the experiment at low energy. Phase shifts due to these potentials are shown in Fig. 9. The number attached to the potentials in Fig. 8 corresponds to the number attached to the curves in Fig. 9. Comparing this with the result of former section, it will be seen that the *S*-wave receives more effects from the hard core than the outer attractive potentials. Then the larger hard core radius makes the phase shift smaller. On the contrary, the phase shifts of waves higher than *P* wave are more effectively affected by the attractive well than the hard core. So the larger radius of the hard core and the deeper attractive well around it give rise to their larger phase shifts. Since the potential of this type have a fairly strong attractive part at meson range or outside of that, the phase shifts of high angular momentum wave, e.g. *F*- or *G*-wave, are larger than those derived from the Yukawa or the exponential potentials.
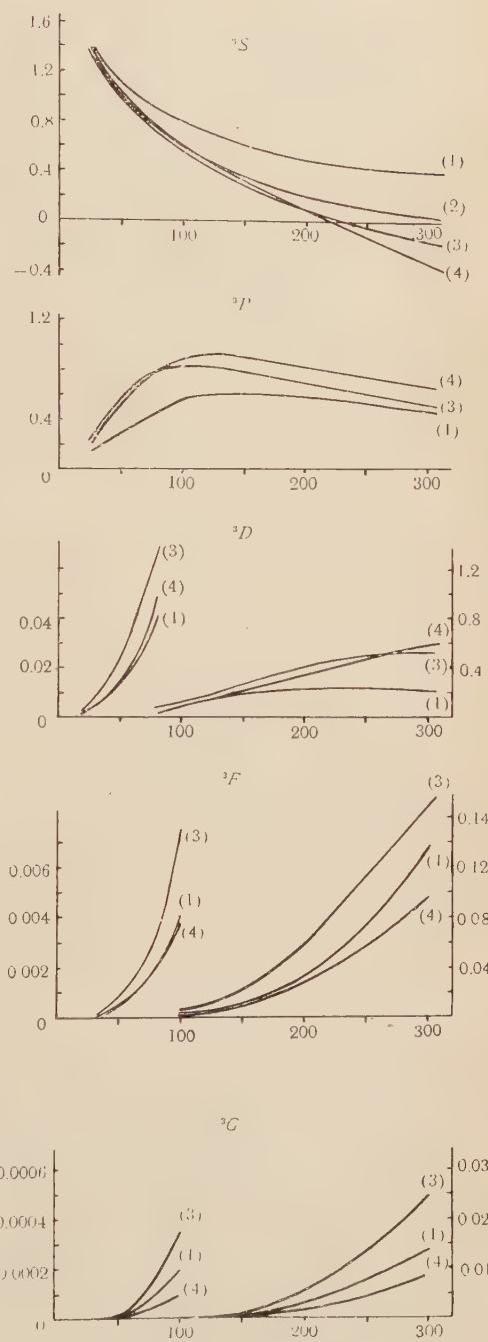
Fig. 9.

The phase shift of each partial wave is plotted as a function of the energy with the potential of Fig. 8. Abscissa : energy ; Ordinate : phase shift.

(b) Next we consider the so called "*L-shape potential*" which consist of the deep well potential (the "body" of the well) and shallow well potential (the "tail" of the well). In Fig. 10 several such potentials are given. Here we assume the same range for both the body and the tail. The potential (4) is the same as that given in Fig. 1. The phase shifts owing to these potentials are shown in Fig. 11. Since there are no remarkable differences between the phase shifts derived from each potential, only phase shifts of S, P, and D wave are plotted. Curves (E) are the phase shift due to the exponential potential adjusted to the low energy experiments. These curves agree well with those due to the potential (4). This result seems to support the standpoint by which any given potential is substituted by the appropriate L-shape potential. Curves (Y) are due to the Yukawa potential. These curves are slightly different from others. But these curves are also reproduced by the L-shape potential which consists of the more narrow and deeper body and more shallow and longer tail.
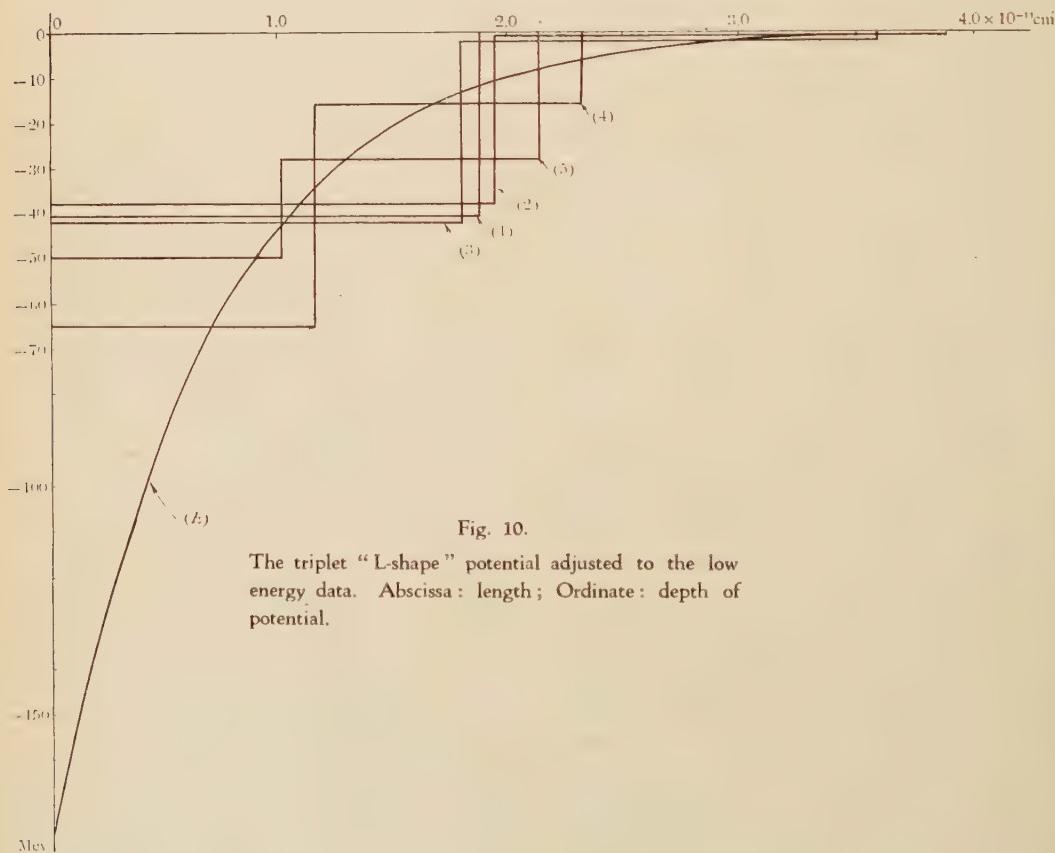
Fig. 10.

The triplet "L-shape" potential adjusted to the low energy data. Abscissa: length; Ordinate: depth of potential.

(ii)   *Singlet potentials*

(a) .Some potentials which consist of *hard core* and a square well are shown in Fig. 12. Of course they are adjusted to the low energy singlet parameters. The phase shifts caused by them are shown in Fig. 13.

(b)   Here, the phase shifts are investigated using each one of the various shapes
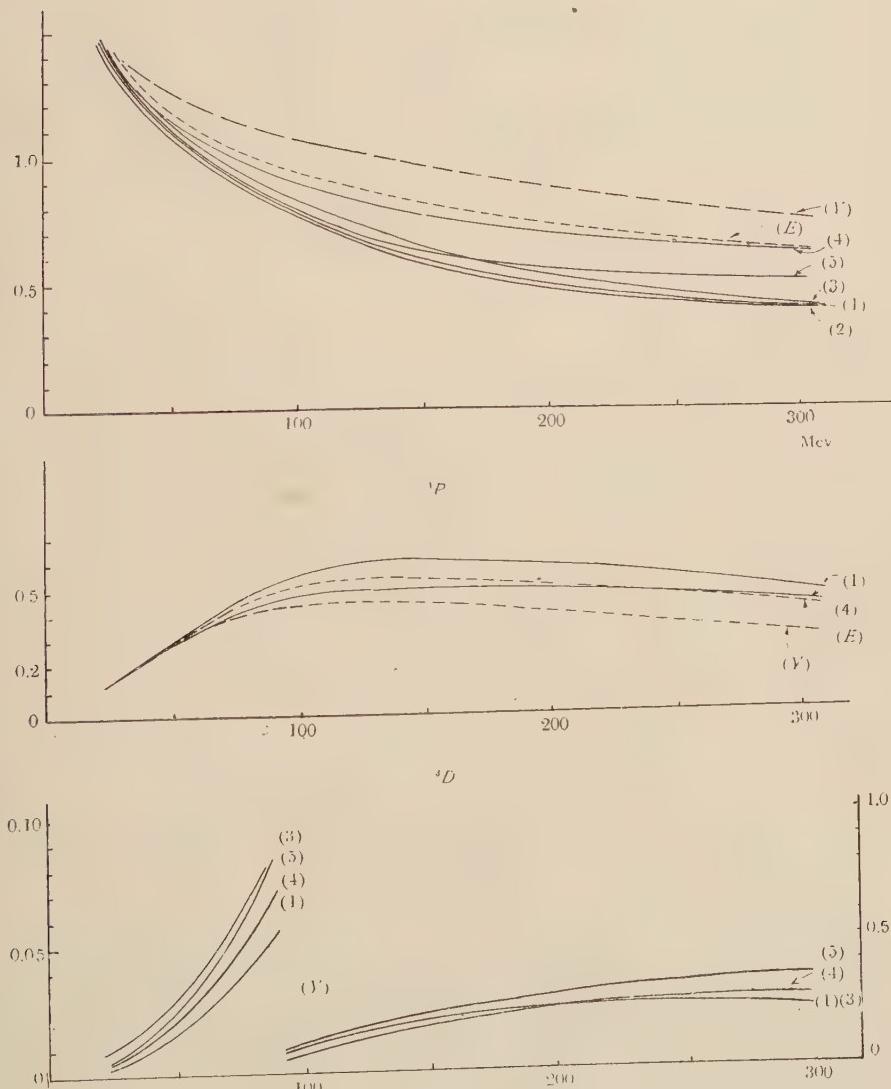
Fig. 11.

The phase shift due to the potential of Fig. 10. .Abscissa : energy ; Ordinate : phase shift.

of attractive well around the hard cores with *nearly fixed radii* as are shown in Fig. 14. The radius of a hard core is about 0.5×10^{-13}cm. There are no remarkable difference at all but S and P-phase shifts. Plots of S-phase shifts versus energy are given in Fig. 15. P phase shift can be roughly estimated from Fig. 4. Potentials (3) and (4) in Fig. 14 are the same as that given in given in Fig. 12 with the same label. As for the higher angular momentum, the phase shifts resulting from the potential (1) agree completely with those from (4), and those from (2) with these from (3). The Jastrow's potential[4] and the potential (3) give nearly the same result. The S-wave phase shift does not affect to the angular distribution so much sensitively, therefore the potential for singlet even state may be approximated by the potential $=$ '*hard core plus square well*'. Conversely it seems to be impossible to determine the potential for the singlet even state more precisely from the scattering experiment.
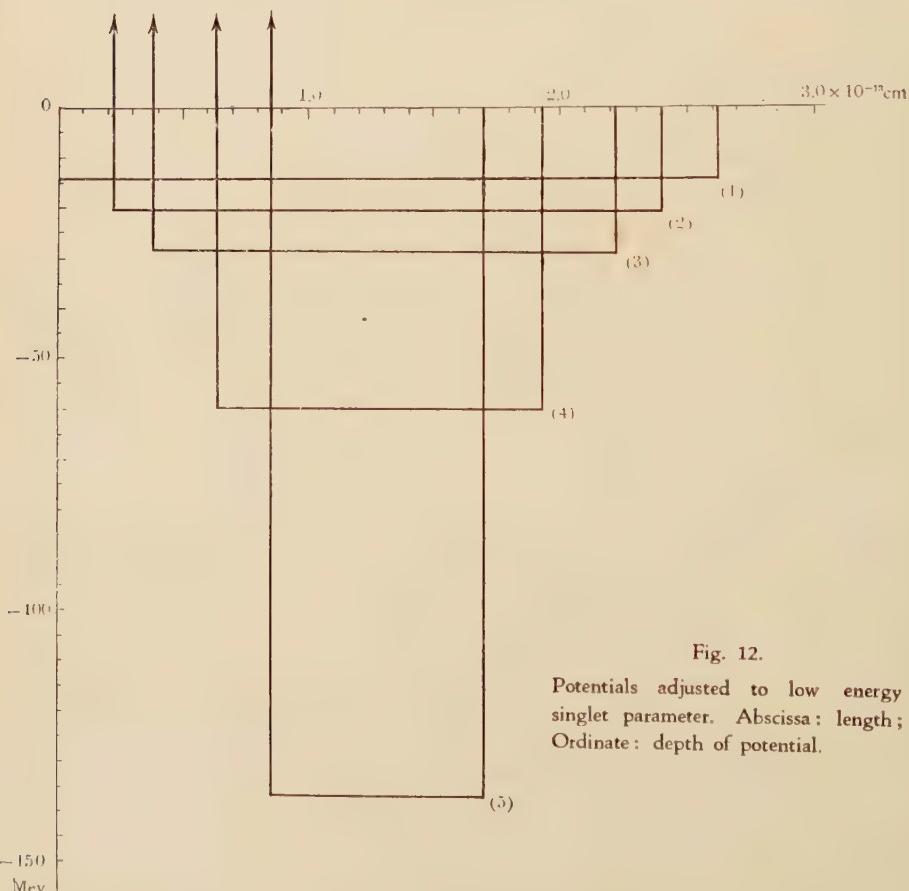
Fig. 12.

Potentials adjusted to low energy singlet parameter. Abscissa: length; Ordinate: depth of potential.

(iii)   *Remark on phase shifts of higher angular momentum*

Using the phase shifts which are obtained in section 2, the effects of the G-wave phase shifts are examined on the angular distribution of the neutron-proton scattering. The angular distribution is given by

$$d\sigma = \text{const.}\left|\sum_l (2l+1)\,e^{i\delta_l}\sin\delta_l\,P_l(\cos\theta)\right|^2 = \text{const.}\sum_{l=0} c_l\,P_l(\cos\theta)$$

where c_l is a function of phase shifts, and the phase shifts and the angular distribution are connected by the coefficients c_l. The c_2/c_0 and c_4/c_0 are calculated with the phase shifts due to the potential in Fig. 1, and the following result is obtained for the incident energy 100 Mev :

$$d\sigma = 5.395[1 + 0.800\,P_2(\cos\theta)$$
$$+ 0.149\,P_4(\cos\theta)]$$

in units millibarn, all c_l/c_0 where $l > 4$ are negligible. If we neglect δ_4, it becomes

$$d\sigma' = 5.395\,[1 + 0.798\,P_2(\cos\theta)$$
$$+ 0.148\,P_5(\cos\theta)]$$

and so the effect of G-wave is negligibly small at this energy. Even in the case of long tailed potential in section 1 (not adjusted to low energy data) the influence of G wave to c_2/c_0 or c_4/c_0 is smaller than 1 per cent. For 200 Mev, these become

$$d\sigma = 2.557\,[1 + 1.937\,P_2(\cos\theta)$$
$$+ 0.608\,P_4(\cos\theta)]$$

$$d\sigma' = 2.555\,[1 + 1.890\,P_2(\cos\theta)$$
$$+ 0.567\,P_4(\cos\theta)]\quad\text{(neglecting }\delta_4)$$

therefore the effect of G wave is appreciable, at this energy c_6/c_0 is also not so small, and so the angular distribution is rather ⊔-shape or a wein bottle shape, i.e., the angular distribution is flat about

Fig. 13.

The phase shift due to the potential in Fig. 12.
Abscissa : energy ; Ordinate : phase shift.

90°, or sometimes appear a small hill about 90°. On the other hand, if experimental results show the ⊔-shape or a wein bottle shape angular distribution, it seems that the G phase shift is fairly large.
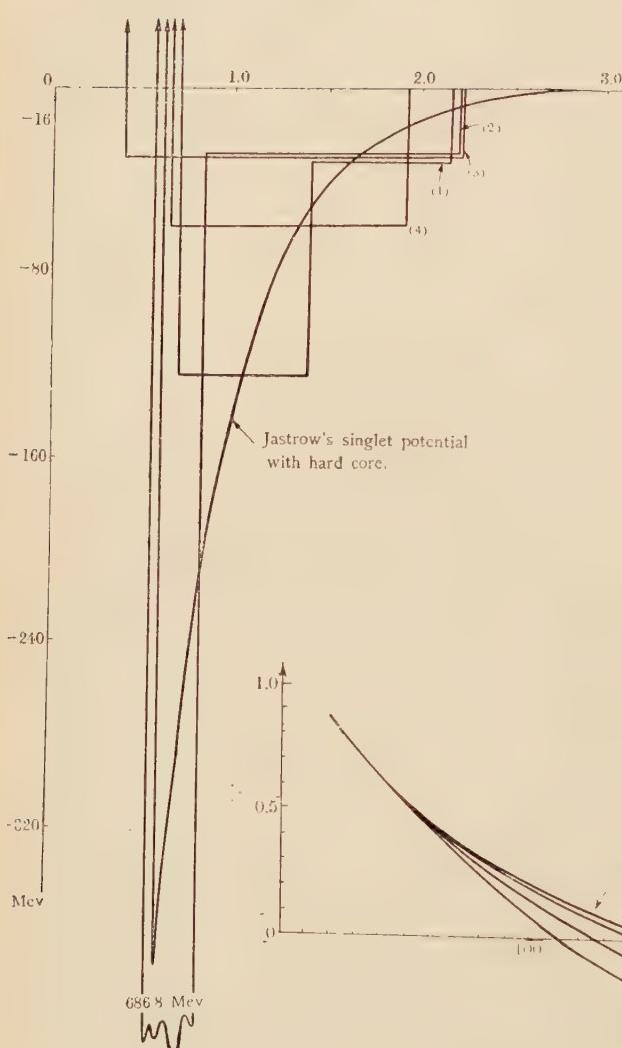
Fig. 14.

Some singlet potentials. Abscissa: length; Ordinate: depth of potential.
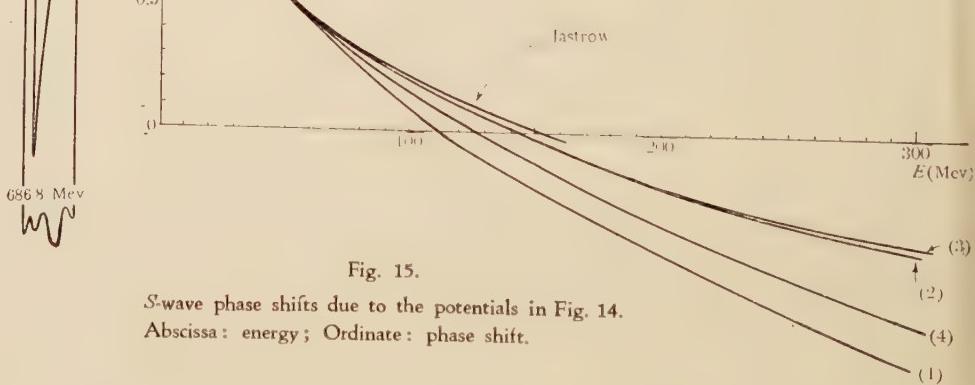
Fig. 15.

S-wave phase shifts due to the potentials in Fig. 14. Abscissa: energy; Ordinate: phase shift.

Conclusional remark

It is seen that S wave (and also P-wave of the incident energy higher than about 100 Mev) are affected by the change of inner part of potentials, but P- and higher angular momentum waves are mainly scattered by the outside region of potentials. Recently Machida and Semba[12] have given the 6th order meson theoretical potential and

shown that the higher order effects in the meson theoretical calculation do not alter the main characters of the already given 2nd and 4-th order potentials in outside region. Therefore, phase shifts of P- or D- wave and higher momentum waves are substantially restricted to the value computed from the meson theoretical potential of 2nd and 4-th order. And we feel much interest in the inside region of the potentials and in the behavior of the S- (and sometimes also of P-) wave phase shifts.

In the present paper, tensor forces are not considered. But the tensor force is very strong in the meson theoretical potential. Moreover, it seems necessary to assume a fairly strong tensor potential in order to explain the double scatttering experiments.[13] Therefore, it is desirable to carry out the similar analysis taking the tensor potential into account. But then the coupled equation such as the deuteron problem must be treated and the procedure will be very troublesome. In the forthcoming paper, the similar problem taking into account of tensor potentials will be treated.

References

1) R. C. Christian and E. W. Hart, Phys. Rev. **77** (1950), 441.
2) R. C. Christian and H. P. Noyes, Phys. Rev. **79** (1950), 85.
3) K. M. Case and A. Pais, Phys. Rev. **80** (1950), 203.
4) R. Jastrow, Phys. Rev. **81** (1951), 165.
5) M. Taketani, S. Nakamura and M. Sasaki, Prog. Theor. Phys. **6** (1951), 581.
6) M. Taketani, S. Machida and S. O-numa, Prog. Theor. Phys. **7** (1952), 46.
7) K. A. Brueckner and K. M. Watson, Phys. Rev. **92** (1953), 1023.
8) D. Bohm and C. Richman, Phys. Rev. **71** (1947), 567.
9) N. F. Mott and H. S. W. Massey, " *The Theory of Atomic Collisions* ". 2nd. ed. (1949), Oxford. p. 25.
10) N. F. Mott and H. S. W. Massey, *loc. cit.* p. 32.
11) M. Kikuta, Soryushiron Kenkyu (Mimeographed circular in Japanese), **4** (1952), No. 8, p. 57.
12) S. Machida and K. Semba, Soryushiron Kenkyu, **5** (1953), No. 11, 1310.
13) L. J. B. Goldfarb and D. Feldman, Phys. Rev. **88** (1952), 1099.

Note on the Second Kind Interaction

Susumu OKUBO

Department of Physics, Tokyo University

The examples which must not be treated by the perturbation are given, and it is shown that there exist some second kind interactions which become to the first kinds, if the uses of the perturbation are avoided. Especially, the pseudoscalar theory with the vector coupling after the Dyson transformation belongs to this, although it was proven only for the vanishing meson mass. At last, the Green functions in the scalar meson theory with the vector coupling are obtained exactly.

§ 1. Introduction

As is well known, since the renormalisation procedure[1] in the quantum electrodynamics has been succeeded in, the same for the other interactions involving mesons was done by many authors.[2] But there exist the interactions where the renormalizations are impossible.

Sakata et al.[3] has classified the any interactions into the first and second kinds and supposed that this classification would be essential in the future theory. Here the first is the renormalisable interaction and the second is not. But as the classification is based on the perturbation expansion, there may be the possibility that the second kind becomes to the first, if the use of the perturbation is avoided. We will show such examples in § 3. One of such examples is the non-linear interaction $\exp g\varphi$, where φ is the scalar photon field operator. As is known from the expansion of the exponential, this is a mixture consisting of the first (ϕ^n with n less than 4) and second interactions ϕ^n with n larger than 5).

So if we treat this according to the perturbation of g, the renormalisations are impossible. But we can show that there are no divergences which are not renormalisable, if we avoid the use of the perturbation of g and the analytic continuation of g^2 is used. This method is based on treating $\exp g^2 D^F$ rightly, by the non-perturbation. This resembles to the Ning Hu[4] and Kamefuchi-Umezawa's[5] in respect to use $\exp g^2 D_F(x)$ as the propagation function instead of $D_F(x)$, whereas they use $D_F'(x)$ instead of $D_F(x)$, after the divergences in $D_F'(x)$ was removed by the appropriate counter-terms. As the matter of fact, $D_F'(x)$ to be taken as the propagation function is not true D_F' but the sum of all graphs which are made of the all graphs repeated by the diagrams up to the some orders of g. This order can be taken arbitrarily. This is a defect of their method and furthermore, in our method there are no needs to subtract any divergence from the propagation function $\exp g^2 D_F(x)$, because it is divergenceless. This method can be extended to the

pseudovector coupling of the pseudoscalar meson after the Dyson transformation. In § 4, the Green functions of the vector coupling of the scalar meson are obtained exactly and studied in the same way.

§ 2. Fourier transformation of $\exp \lambda D_F(x)$

In this chapter, we show that there are examples which have no divergences if the perturbations are not used. Put

$$f(x)=1/(u+D_F(x)) \tag{1}$$

or

$$f(x)=1/u(1-1/u\,D_F(x)+1/u^2\,D_F{}^2(x)\cdots\cdots). \tag{1}'$$

It is clear that the Fourier transformations of every terms except the first and second of $(1)'$ diverge. From (1) or $(1)'$

$$f(x)=1/u-1/u\,D_F(x)\,f(x).$$

Put

$$f(p)=\int\,\exp\,(ipx)f(x)d^4x \tag{2}$$

and notice that*

$$D^F(x)=\frac{-i}{(2\pi)^4}\int\frac{1}{p^2-i\epsilon}\exp(ipx)d^4\,p$$

then

$$f(p)=\frac{(2\pi)^4}{u}\,\delta(p)+\frac{i}{(2\pi)^4u}\int\frac{1}{(p-p')^2}\,f(p')d^4p',$$

or by putting

$$f(p)=\frac{(2\pi)^4}{u}\,\delta(p)+g(p) \tag{3}$$

$$g(p)=\frac{i}{u^2}\,\frac{1}{p^2}+\frac{i}{(2\pi)^4u}\int\frac{1}{(p-p')^2}\,g(p')d^4p'. \tag{4}$$

Following to Edwards' method,[6] we suppose this solution is given by

$$g(p)=\int_0^\infty\frac{F(t)\,dt}{(p^2+16\pi^2ut)^3}. \tag{5}$$

Here F is the function to be determined after and p^2 in the denominator of the right hand of (5) has a vanishingly small negative imaginary part implicitly. Then by the formula (see Appendix A)

$$\int\frac{1}{(p-p')^2}\,\frac{1}{(p'+\lambda)^3}\,d^4\,p'=\frac{i\pi^2}{2}\,\frac{1}{\lambda(p^2+\lambda)} \tag{6}$$

---

*) Our definitions of D_F is the same as in Gell-Mann and Low. Other symbols after this are of the same authors',

(4) becomes to

$$\int_0^\infty \frac{F(t)}{(p^2+16\pi^2 ut)^3}\, dt = \frac{i}{u^2}\frac{1}{p^2} - \frac{1}{2(16\pi^2 u)^2}\int_0^\infty \frac{F(t)}{t(p^2+16\pi^2 ut)}\, dt. \qquad (4)'$$

Here u must be the number which is real or with the vanishingly small negative imaginary part implicitly, in order that the formula (6) can be used. Notice that

$$1/(p^2+16\pi^2 ut)^3 = 1/2(16\pi^2 u)^2\, d^2(p^2+16\pi^2 ut)^{-1}/dt^2$$

and perform the partial integration of the left hand of (4)' then (4)' is satisfied if the next equations hold good.

$$d^2F/dt^2 + F/t = 0, \quad (F/t^2)_\infty = (F'/t)_\infty = 0, \quad F(0)=0, \quad F'(0)=2i(16\pi^2)^2. \quad (7)$$

The function F which satisfies (7) is given by

$$F(t) = 2i(16\pi^2)^2 \sqrt{t}\, J_1(2\sqrt{t}). \qquad (7)'$$

Therefore from (5)

$$g(p) = \frac{2i}{\pi^2 u^3}\int_0^\infty \frac{t^2 J_1(t)}{(t^2+p^2/4\pi^2 u)^3}\, dt. \qquad (5)'$$

By the use of the formula[7]

$$\int_0^\infty \frac{x^{\nu+1}}{(x^2+k^2)^{\mu+1}} J_\nu(ax)\, dx = \frac{a^\mu k^{\nu-\mu}}{2^\mu \Gamma(\mu+1)} K_{\nu-\mu}(ak),$$

we get

$$g(p) = \frac{i}{4\pi^2 u^3}\frac{1}{\sqrt{p^2/4\pi^2 u}} K_1(\sqrt{p^2/4\pi^2 u}). \qquad (8)$$

Here

$$K_1(x) = \left(\log\frac{x}{2}+\gamma\right)\sum_{n=0}^\infty \frac{x^{1+2n}}{2^{1+2n}\, n!(n+1)!} + \frac{1}{x} - \frac{x}{4} - \sum_{n=1}^\infty \frac{x^{1+2n}}{4^{1+n} n!\,(n+1)!}$$

$$\left\{\sum_{p=1}^n \left(\frac{1}{p}+\frac{1}{p+1}\right)+1\right\}.$$

Therefore $g(p)$ contains the logarithmic terms of $1/u$ by (8). This shows that the perturbation by $1/u$ is impossible and so divergences arise if the perturbations are forced. Namely these divergences result from the insufficient mathematical method and are not essential. But it can be seen that the parts which are not divergent in the fourier transformation of (1)' are identical to the corresponding terms of the expansion of (8).
    Nextly let us consider the fourier transformation of the function

$$I(x) = \exp(-i\lambda D_F(x)) \quad (\lambda>0). \qquad (9)$$

Each term except the first and second is also divergent if we expand (9) by the power series of λ and take the fourier transformation. Notice that $D_F(x)$ has the negative imaginary part when it exists, by
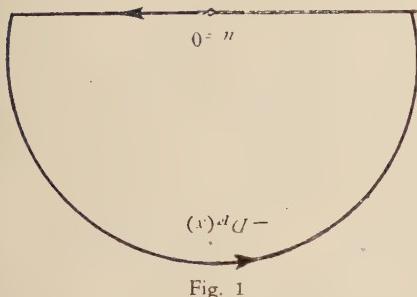
Fig. 1

$$D_F(x) = \frac{1}{4\pi^2} \frac{1}{x^2 + i\epsilon}$$

$$= \frac{1}{4\pi^2}\left(\frac{1}{x^2} - i\pi\,\delta(x^2)\right) \tag{10}$$

Then as is known from Fig. 1,

$$I(x) = \frac{1}{2\pi i}\int_{-\infty}^{\infty}\frac{1}{u + D_F(x)}\exp(iu\lambda)\,du . \tag{9}'$$

Here u in the denominator must have the vanishingly small negative imaginary part. Then by using (3) and (5)' and integrating on u, we get

$$I(p) = \frac{1}{2\pi i}\int_{-\infty}^{\infty}f(p)\exp(iu\lambda)\,du = (2\pi)^4\delta(p) - \frac{\lambda^2 i}{\pi^2}\int_0^\rho dt\,\frac{J_1(t)}{t^4}\exp\left(-i\,\frac{p^2\lambda}{4\pi^2 t^2}\right). \tag{11}$$

Or, by integrating partially and noticing that p^2 has a vanishingly small negative imaginary part

$$I(p) = (2\pi)^4\delta(p) - \frac{2\lambda}{p^2}\int_0^\infty dt\,\frac{J_2(t)}{t}\exp\left(-i\,\frac{p^2\lambda}{4\pi^2 t^2}\right) \tag{11'}$$

Returning to the coordinate space and using (11)

$$I(x) = 1 - \int_0^\infty dt\,J_1(t)\exp\left(i\,\frac{\pi^2 t^2}{\lambda}x^2\right). \tag{12}$$

It is easily shown that (12) is identical to (9) by noticing that x^2 in (12) contains a vanishingly small positive imaginary part and expanding $J_1(t)$ in the power series of t.

Let put as follows for brevity

$$\varphi(u) = u\int_0^\infty dt\,\frac{J_2(t)}{t}\exp\left(i\,\frac{u}{t^2}\right) \qquad (u = -p^2\lambda/4\pi^2) \tag{13}$$

and then

$$I(p) = (2\pi)^4\delta(p) + 8\pi^2/(p^2)^2\cdot\varphi(u). \tag{13}'$$

By the direct calculation, it is seen $\varphi(u)$ satisfies the next third order differential equation

$$\left(\frac{d^3}{du^3} - \frac{1}{4iu^2}\right)\varphi = 0. \tag{14}$$

The solutions of the differential equation

$$(u^2\,d^3/du^3 - 1)w = 0 \tag{15}$$

are given by

$$w_1(x) = \sum_{n=0}^{\infty} \frac{1}{(n+2)!\,(n+1)!\,n!}\, x^{n+2}$$

$$w_2(x) = w_1(x)\log x + x + \sum_{n=0}^{\infty} \frac{x^{n+2}}{(n+2)!\,(n+1)!\,n!}\left\{ 3\gamma - \sum_{S=1}^{n+2}\frac{1}{S} - \sum_{S=1}^{n+1}\frac{1}{S} - \sum_{S=1}^{n}\frac{1}{S}\right\}$$

$$w_3(x) = \log x\,(2w_2 - w_1 \log x) - 2 + 2(3\gamma - 1)x + \cdots\cdots .$$

$$(15)'$$

From (13), $\varphi(0) = 0$. So φ is a linear combination of w_1 and w_2. In fact as will be shown in the Appendix B,

$$\varphi(u) = \frac{u}{8}\int_{c-i\infty}^{c+i\infty} dt\,\frac{1}{t^2}\,(\exp t)\,\sqrt{\frac{iu}{t}}\,H_1^{(1)}\!\left(\sqrt{\frac{iu}{t}}\right)\ (u>0)$$

$$= -\frac{u}{8}\int_{c-i\infty}^{c+i\infty} dt\,\frac{1}{t^2}\,(\exp t)\,\sqrt{\frac{iu}{t}}\,H_1^{(2)}\!\left(\sqrt{\frac{iu}{t}}\right)\ (u<0)$$

$$(16)$$

or

$$\varphi(u) = 2iw_2(-ui/4)\quad (u>0)$$

$$\varphi(u) = 2iw_2(-ui/4) - 4\pi i w_1(-ui/4)\ (u<0).$$

$$(17)$$

Here argument of u must be taken as

$$|\arg\sqrt{iu/t}| < \pi/2 .$$

From (15)', (17) and (13), $I(p)$ has the logarithmic terms of λ. So the perturbation by λ is meaningless as before, and the divergences give arise for this, if it is done.

The behaviour of $I(p)$ when p^2 becomes to the infinity, is given by

$$I(p) = (2\pi)^4 \delta(p) + \text{const}\cdot(p^2)^{-4/3}\exp\{3i(-p^2\lambda/16\pi^2)^{1/3}\}.$$

$$(18)$$

This will be shown in the Appendix C.

At last we consider the case of Δ_F instead of D_F.

Put

$$f(x) = 1/(u + \Delta_F(x))$$

$$(19)$$

then

$$f(p) = \frac{(2\pi)^4}{u}\delta(p) + \frac{i}{(2\pi)^4 u}\int\frac{1}{(p-p')^2 + m^2}f(p')d^4p'$$

$$= \frac{(2\pi)^4}{u}A(p) + \frac{i}{(2\pi)^4 u}\int\frac{1}{(p-p')^2}f(p')d^4p'.$$

$$(20)$$

Here

$$A(p) = \delta(p) - \frac{im^2}{(2\pi)^8}\int\frac{1}{[(p-p')^2 + m^2](p-p')^2}f(p')d^4p'\quad(20)'$$

The solution of (20) is

$$f(p) = \int f_0(p-p')A(p')d^4p',$$

where $f_0(p)$ is the function given by $(5)'$ and (3), which is written merely as $f(p)$ before.

Then by $(20)'$

$$f(p) = f_0(p) - \frac{im^2}{(2\pi)^3} \int dp' dp'' f_0(p-p') \frac{1}{[(p'-p'')^2 + m^2](p'-p'')^2} f(p''). \quad (21)$$

This is solved by the interation method. Each terms are divergenceless.

3. Analysis of divergence of some interactions

In this chapter we consider the interaction $\exp g\phi$ at first. We take the meson mass, for zero, for it will be complicated as will be seen from last discussion of the chapter II. But this limitation will be not essential for the analysis of the divergences.

As is stated in the introduction, this interaction is a mixture of the first and second kinds which are ϕ^n with n less than 4 and ϕ^n with n larger than 5, respectively. As the constant term in the Hamiltonian changes only the constant phase of S matrix, we use the following expression as the interaction.

$$\left. \begin{array}{l} H = f \exp(g\phi) - f^* = f\{\exp(g\phi) - \exp(g^2 D_F(0)/2)\} \\ f^* = f \exp(g^2 D_F(0)/2). \end{array} \right\} \quad (22)$$

Here f^* represents the renormalised coupling constant.
S matrix is given by

$$S = P\left\{ \exp\left(-i \int H d^4 x\right) \right\} = \sum_{n=0}^{\infty} \frac{(-if)^n}{n!} \int$$

$$\cdots \int dx_1 \cdots dx_n P\left\{ \prod_{i=1}^{n} (\exp g\phi(x_i)) - \exp\left(\frac{g^2}{2} D_F(0)\right) \right\}. \quad (23)$$

Following to the method given by Glauber[8]

$$P\left\{ \prod_{i=1}^{n} \exp g\phi(x_i) \right\} = \exp\left\{ n \frac{g^2}{2} D_F(0) \right\} \prod_{i>j}^{n} \exp g^2 D_F(x_i - x_j) : \prod_{i=1}^{n} \exp g\phi(x_i) :. \quad (24)$$

Here $: A :$ means Wick's[9] ordered operator of A.
Now let us consider the case $n = 2$, as a example. Using (24), the integrand of the second order term of (23) becomes to

$$f^2 P\{(\exp g\phi(x_1) - \exp g^2/2 D_F(0))(\exp g\phi(x_2) - \exp g^2 D_F(0))\}$$

$$= (f^*)^2 [\exp g^2 D_F(x_1 - x_2) \exp g\phi(x_1) \exp g\phi(x_2) - \exp g\phi(x_1) - \exp g\phi(x_2) + 1]. \quad (25)$$

For brevity, the symbols of the ordered operators are dropped. $\exp g^2 D_F(x_1 - x_2)$ corresponds to the propagation between x_1 and x_2, but it contains the constant number 1 according to the expansion of the exponential. As this has not to deal with the propagation,

it has to be eliminated. By the same reason, the factor which represents the external line is not $\exp g\phi(x)$ but $\exp g\phi(x)-1$. Then (25) is rewritten as follows:

$$(f^*)^2 [(\exp g^2 D_F(x_1-x_2)-1)(\exp g\phi(x_1)-1)(\exp g\phi(x_2)-1)$$

$$+ (\exp g^2 D_F(x_1-x_2)-1)(\exp g\phi(x_1)-1) + (\exp g^2 D_F(x_1-x_2)-1)(\exp g\phi(x_2)-1)$$

$$+ (\exp g^2 D_F(x_1-x_2)-1) + (\exp g\phi(x_1)-1)(\exp g\phi(x_2)-1). \qquad (25)'$$

Each term of (25)' corresponds to each graph given in the Fig. 2. Here the factor $\exp g^2 D_F(x_1-x_2)-1$ is corresponded to the internal line which runs between x_1 and x_2, and the factor $\exp g\phi(x)-1$ to the external line which departs externally from x. As is known from the expansion of the exponential, the more external lines than one can radiate from the one vertice. It is to be noticed that there is not such a vertice that has neither the the internal nor external line which depart from it. This is the result by the subtraction of the constant term from the Hamiltonian.
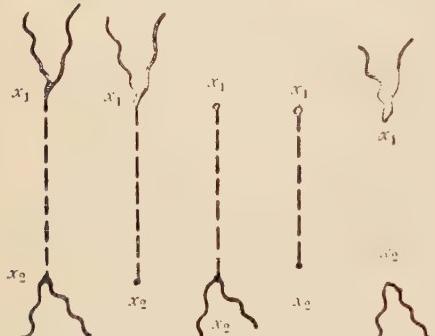
Fig. 2

The rules mentioned above are generalized in the arbitary n vertice graph. Namely the rules are given as follows.

(A)   Between the two arbitrary vertices x_1 and x_2, the internal line may be taken or not. If taken, the factor $\exp g^2 D_F(x_1-x_2)-1$ is given to it. There are no limitations on the number of the internal lines which depart from the one vertice.

(B)   From the each vertice, the external lines may be radiated or not, If radiated, the factor $g^m\phi^m$ is given to it. Here m is the number of the external lines. And there are not such vertices that have neither the internal nor external lines which depart from it.

(C)   At last the factor $(-if^*)^n/n!$ is multiplied.

After the preparations above, we will show there are no divergences more, when $\exp g^2 D_F(x_1-x_2)$ is not expanded to the power series of g^2 and the analytic continuation method is used for g^2.

We suppose that g^2 is a purely negative imaginary.

$$g^2 = -i\lambda \quad (\lambda > 0)$$

Then the factor corresponding to the internal line which runs between x_1 and x_2 is given by (12) as follows:

$$\exp(-i\lambda D_F(x_1-x_2))-1 = -\int_0^\infty dt\, J_1(t) \exp\left\{ i\frac{\pi^2 t^2}{\lambda}(x_1-x_2)^2 \right\}. \qquad (27)$$

For the analysis of the divergences, this form is more convenient than (11) or (11)' in the momentum representation. The integral variable t in (27) corresponds to the momentum of the line. Such variable is named as the internal variable.

Let us consider a n vertice graph and integrate over the n coordinate variables at first. The last of this n integrations gives only the factor $(2\pi)^4 \delta(\varepsilon_i p_i)$ which means the total energy momentum conservation for the external lines. Integrations on the other $(n-1)$ variables give a fraction of the order $4(n-1)$ of the internal variables according to the equation

$$\int d^4x \exp(i a^2 x^2) = i\pi^2/a^4. \tag{28}$$

The ultra-violet divergences correspond to the large a.

And the integrals over all internal variables give the convergent result, because the integrand of the integral over a internal variable t contains the factor $J_1(t)$ by (27), and the next integrals converge necessarily for the large t.

$$\int_0^\infty \frac{J_1(t)}{(t+c)^n} \, dt \quad \text{(for } n \geq 0).$$

So, there do not give rise any divergences.

Thus, although we have not performed the calculations by the power series expansions of g, we have done it for f^*. As is shown by Hurst and others[10] the number of graphs for the given n is so large even for the interaction $f\phi^3$, that the perturbation series of f will not converge. In the case treated now, the number is much larger than in ϕ^3, and soo ur series will not converge. It is desirable to find the method for the treatment by the non-perturbation method for f.

Nextly let us consider the interactions of the form $M\bar{\psi}(\exp g\phi)\psi$ and $M\bar{\psi}(\exp i g\gamma_5\phi)\psi$. Here M is the nucleon rest mass. Especially, the latter is the pseudovector coupling of the neutral pseudoscalar meson with the nucleon, transformed by the Dyson tranformation.

In this case, the meson mass is put to be zero for brevity. The nucleon mass term is included in the free Hamiltonian, and so the interaction term becomes to

$$\left. \begin{array}{c} M\bar{\psi}(\exp g\phi - \exp g^2 D_F(0)/2)\psi \\[2mm] M\bar{\psi}\{\exp i g\gamma_5\phi - \exp(-g^2 D_F(0)/2)\}\psi. \end{array} \right\} \tag{29}$$

or

The renormalised mass is

$$\left. \begin{array}{c} M^* = M\exp(g^2 D_F(0)/2) \quad \text{(scalar case)} \\[2mm] M^* = M\exp(-g^2 D_F(0)/2) \quad \text{(pseudoscalar case)}. \end{array} \right\} \tag{29'}$$

or

The graphical construction rules for the S matrix are the same to the before, except that the nucleon line must be taken, and for the pseudoscalar case, the factor given to the internal meson line must be given by

$$\exp(-g^2\gamma_5(1)\gamma_5(2)D_F(1-2)) - 1 = 1/2 \cdot (1 - \gamma_5(1)\gamma_5(2))(\exp g^2 D_F(1-2) - 1)$$
$$+ 1/2 \cdot (1 + \gamma_5(1)\gamma_5(2))(\exp(-g^2 D_F(1-2)) - 1). \tag{30}$$

Again the analytic continuation method for g^2 is used. For the scalar case, g^2 is put to

$-i\lambda(\lambda > 0)$ as before, but for the pseudoscalar case g^2 in the factor $\exp(g^2 D_F(1-2))$ is put to $-i\lambda$ and g^2 in the factor $\exp(-g^2 D_F(1-2))$ to $+i\mu(\mu > 0)$. Then for the internal meson, the equation (27) can be used. For the internal nucleon line, the next is used

$$S_F(x) = \frac{i}{2\pi^2} \int_0^\infty (2i a^3 \gamma_\mu x_\mu - Ma) \exp\left(i a^2 x^2 - \frac{iM^2}{4a^2}\right) da. \tag{31}$$

For the analysis, the circumstances are almost the same, and so the analysis is done for the scalar case.

As before let us consider a n vertice graph and perform the integrations over the n coordinate variables at first. Then differing from the before, the integrals of the form give arise

$$\int d^4x \; x^2 \exp(i a^2 x^2). \tag{32}$$

Though these integrals are divergent at first sight, this is not the true difficulty. These integrals appear in the case of the quantum electrodynamics, if the integral representations of the form (31) are used for $S_F(x)$ and $D_F(x)$. The right method therein is to interpret (32) as follows:

$$\int d^4x x^2 \exp(i a^2 x^2) = \frac{1}{i} \frac{d}{d(a^2)} \int d^4x \exp(i a^2 x^2) = -\frac{2\pi^2}{a^6}. \tag{32'}$$

Then, suppose that the graph contains the F_N internal lines of the nucleons. By the integrations over all coordinate variables, the order of the internal variables appearing in the denominator of the resulting fraction is given by

$$4(n-1) - 2F_N. \tag{33}$$

Nextly we perform the integrations over all nucleon internal variables. The condition for the convergence is clearly

$$(4n-1) - 3F_N \geq 1. \tag{33'}$$

As the integrations over the meson variables converge necessarily by the same reason given before, this gives the condition for the convergence. If the number of the external nucleon lines are E_N,

$$F_N = n - 1/2 E_N.$$

Therefore (33)' becomes to

$$n + 3/2 E_N \geq 5. \tag{33''}$$

The cases where (33)'' is not satisfied are

$$E_N = 0, \; n = 2, 3, 4 \tag{34}$$

Fig. 3

The examples of the case are the graphs given in Fig. 3. For the subtractions of these divergences it is convenient to use the momentum representations (11) or (11)′ for the lines which depart externally from these closed loops. Then the forms of the divergent integrals are

$$\int_0^\infty dt\, R(t)\exp\left(-i\frac{p^2}{t^2}\right). \quad (35)$$
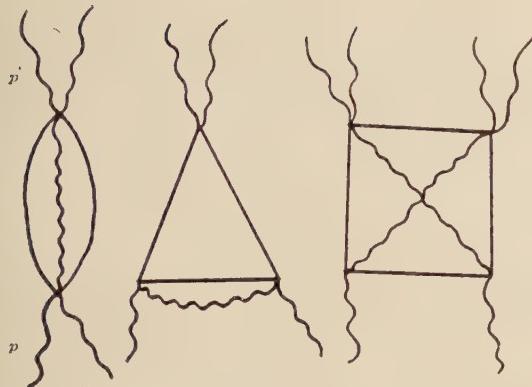
The reason of the appearance of the exponential factor in (35) will be made clear by the next example of the calculation.

$$\int_0^\infty dt\int d^4x\,\exp(ipx)\exp(it^2x^2)=\int_0^\infty dt\,\exp\left(-i\frac{p}{4t^2}\right)\int d^4x\,\exp\left[it^2\left(x+\frac{p}{2t^2}\right)^2\right]$$

$$=i\pi^2\int_0^\infty dt\,\frac{1}{t^4}\,\exp\left(-i\frac{p^2}{4t^2}\right).$$

The separation of the divergence is done as follows

$$\int_0^\infty dt\, R(t)\left[\exp\left(-i\frac{p^2}{t^2}\right)-1\right]+\int_0^\infty dt\, R(t). \quad (35)′$$

The second integral is divergent and the order of the integrand of the first for the large t is increased by two. Therefore for the graph with $n=3$ or 4 this term converges. For the graph with $n=2$, the divergence remains yet. The divergence is eliminated further as follows :

$$\int_0^\infty dt\, R(t)\left(\exp\left(-i\frac{p^2}{t^2}\right)-1+i\frac{p^2}{t^2}\right)+\int_0^\infty dt\, R(t)-ip^2\int_0^\infty dt\,\frac{R(t)}{t^2}. \quad (35)″$$

Thus all divergences are eliminated.
As is known from the forms of the equations (35)′ and (35)″, these infinite quantities can be cancelled, if the initial Lagrangean have the additional terms given by

$$\mathcal{L}_1=\sum_{n=2}^4 A_n(\exp g\phi-\exp g^2 D_F(0)/2)^n+B(\partial \exp g\phi/\partial x_\mu)^2. \quad (36)$$

Here A_n and B are the divergent constants.
So there are no more divergences.

4. Green functions in the vector coupling theory of the scalar meson

It is shown by Dyson[11] that the neutral scalar meson theory with the vector coupling

is trivial, because the Hamiltonian becomes to the free by the Dyson transformation. But the Green functions have the values differing from the free, and they are divergent if the perturbation is used. Here we will obtain the exact formula for the Green functions and show these divergences will be trivial if the perturbation will be avoided.

We depart from the Schroedinger representation.

$$\left.\begin{aligned}
H &= H_0 + H_1, \\
H_0 &= \frac{1}{2}\left[\int \psi^* \boldsymbol{\alpha p}\, \psi\, dV - \int \boldsymbol{\alpha p}\psi\psi^* dV\right] + \frac{m}{2}\int (\psi^*\psi - \psi\psi^*)dV \\
&\quad + \frac{1}{2}\int (\pi^2 + (\nabla\phi)^2 + \mu^2\phi^2)dV, \\
H_1 &= \frac{g}{2}\left[\int \psi^*(\pi + \alpha\nabla\phi)\psi dV - \int (\pi + \alpha\nabla\phi)\psi\psi^* dV\right] + \frac{g^2}{8}\int (\psi^*\psi - \psi\psi^*)^2 dV.
\end{aligned}\right\}$$

$$(37)$$

This Hamiltonian is transformed as follows and becomes to the free.

$$S = \frac{g}{2}\int (\psi^*\psi - \psi\psi^*)\phi\, dV, \tag{38}$$

$$H' = \exp(iS) H \exp(-iS) \equiv H_0. \tag{38}''$$

Therefore the true vacuum state Ψ_0 is connected to the free vacuum Φ_0, by the relation

$$\Psi_0 = \exp(-iS)\,\Phi_0. \tag{39}$$

We will write the operators corresponding to the Heisenberg, interaction and Schroedinger representations by ψ_H, ψ_{int} and ψ respectively.
Then

$$\left.\begin{aligned}
\psi_H(x) &= \exp(iHt)\psi(\boldsymbol{x})\exp(-iHt) \\
\psi_{int}(x) &= \exp(iH_0 t)\psi(\boldsymbol{x})\exp(-iH_0 t).
\end{aligned}\right\} \tag{40}$$

So

$$\exp(iS)\psi_H(x)\exp(-iS) = \exp(iH_0 t)\exp(iS)\psi(\boldsymbol{x})\exp(-iS)\exp(-iH_0 t).$$

And noticing that[12)]

$$\exp(iS)\psi(\boldsymbol{x})\,\exp(-iS) = \exp(-ig\phi(\boldsymbol{x}))\psi(\boldsymbol{x}).$$

Therefore

$$\exp(iS)\psi_H(x)\exp(-iS) = \exp(-ig\phi_{int}(x))\psi_{int}(x). \tag{41}$$

The one body Green function is defined by

$$G(x, y) = i(\Psi_0, T(\psi_H(x)\bar\psi_H(y))\Psi_0).$$

From (39) and (41)

$$G(x,y)=i(\varPhi_0,\ T(\psi_{int}(x)\bar{\psi}_{int}(y)\exp(-ig\phi_{int}(x))\exp(+ig\phi_{int}(y)))\varPhi_0).$$

Following Glauber,[8] this becomes to

$$G(x,y)=iS_F(x-y)\exp g^2(\varDelta_F(x-y)-\varDelta_F(0)).$$

By the same method the two body Green function is given by

$$G(x,x'\ ;\ y,y')=-(\varPsi_0,\ T(\psi_H(x)\psi_H(x'),\ \bar{\psi}_H(y')\bar{\psi}_H(y'))\varPsi_0)$$

$$=\{S_F(x-y)S_F(x'-y')-S_F(x-y')S_F(x'-y)\}\exp(-2g^2\varDelta_F(0))$$

$$\times\exp g^2[\varDelta_F(x-y)+\varDelta_F(x-y')+\varDelta_F(x'-y)+\varDelta_F(x'-y')-\varDelta_F(x-x')-\varDelta_F(y-y')].$$

$$(43)$$

If there is the additional external source J_μ, the Green function is given by the following instead of (42).

$$G(x,y)=iS_F(x-y)$$

$$\times\exp\left[-g\int(\varDelta_F(x-x')-\varDelta_F(y-x'))\frac{\partial J_\mu'}{\partial x'_\mu}d^4x'+g^2(\varDelta_F(x-y)-\varDelta_F(0))\right].\ (44)$$

This can be proven as follows, too. Following Schwinger,[13] the Green function satisfies the next equation

$$(\gamma_\mu\partial_\mu+M+ig\gamma_\mu\langle\partial\phi/\partial x_\mu\rangle+g\gamma_\mu\delta/\partial J_\mu(x))G(x,y)=\delta(x-y).\qquad(45)$$

On the other hand from the equation of motion of the meson

$$(\Box-\mu^2)\langle\phi\rangle=\partial J_\mu/\partial x_\mu.$$

This solution is

$$\langle\phi\rangle=-i\int\varDelta_F(x-x')\frac{\partial J_\mu'}{\partial x_\mu'}d^4x'.\qquad(46)$$

It is easily shown that the solution of (45) and (46) is (44).

    The Fourier transformation of the Green function (42) or (43) are divergent if the perturbation for g^2 is used. But it is clear by the analysis of § 3 that these converge if the analytic continuation is made and the meson mass is put to be zero. The divergent term is only the trivial multiplicative factor $\exp(-g^2\varDelta_F(0))$ which is the amplitude renormalisation constant.

Appendix A

This is ascertained by the direct calculation

$$I=\int\frac{1}{(p-p')^2(p'^2+\lambda)^3}d^4p'$$

Using the relation

$$\frac{1}{ab^3} = 3 \int_0^1 du \frac{(1-u)^2}{[au+b(1-u)]^4}$$

$$I = 3 \int_0^1 du (1-u)^2 \int d^4p' \frac{1}{[p'^2 + p^2 u(1-u) + \lambda(1-u)]^4},$$

By the equation

$$\int d^4p' \frac{1}{(p'^2 + \lambda)^4} = \frac{i\pi^2}{6} \frac{1}{\lambda^2}$$

Finally we get

$$I = \frac{i\pi^2}{2} \int_0^1 \frac{du}{(p^2 u + \lambda)^2} = \frac{i\pi^2}{2} \frac{1}{\lambda(p^2+\lambda)}.$$

This is the equation (6).

Appendix B

$$\varphi(u) = u \int_0^\infty \frac{J_2(x)}{x} \exp\left(i \frac{u}{x^2} \right) dx .$$ (13)

By the formula of Watson's Bessel Function p. 177

$$J_2(x) = \frac{1}{2\pi i} \frac{x^2}{4} \int_{c-i\infty}^{c+i\infty} dt \frac{1}{t^3} \exp\left(t - \frac{x^2}{4t} \right) \quad (c > 0)$$

so

$$\varphi(u) = \frac{1}{2\pi i} \frac{u}{4} \int_{c-i\infty}^{c+i\infty} dt \frac{\exp t}{t^3} \int_0^\infty dx\, x \exp\left(-\frac{x^2}{4t^2} + i \frac{u}{x^2} \right).$$

Putting

$$y = (4itu)^{1/2}/x^2,$$

$$\varphi(u) = \frac{1}{2\pi i} \frac{u}{4} \int_{c-i\infty}^{c+i\infty} dt \frac{1}{t^2} \left(\frac{iu}{t} \right)^{1/2} \exp t \int_0^{(itu)^{1/2}\infty} dy \frac{1}{y^2} \exp \frac{1}{2} \left(\frac{iu}{t} \right)^{1/2} \left(y - \frac{1}{y} \right).$$

Here the argument is taken as it satisfies the condition

$$| \arg(itu)^{1/2} | < \frac{\pi}{2}, \quad \left| \arg\left(\frac{iu}{t} \right)^{1/2} \right| < \frac{\pi}{2} .$$

If $u > 0$, the integration path C_1 is transformed to the path C given by Fig. 1.
    By the formula of Watson's Bessel Function p. 178

$$\int_c dy \frac{1}{y^{n+1}} \exp \frac{z}{2} \left(y - \frac{1}{y} \right) = \pi i H_n^{(1)}(z) \quad | \arg z | < \frac{\pi}{2}$$

so,

$$\varphi(u) = \frac{u}{8} \int_{c-i\infty}^{c+i\infty} dt \frac{1}{t^2} \left(\frac{iu}{t} \right)^{1/2} H_1^{(1)} \left(\sqrt{\frac{iu}{t}} \right) \exp t .$$
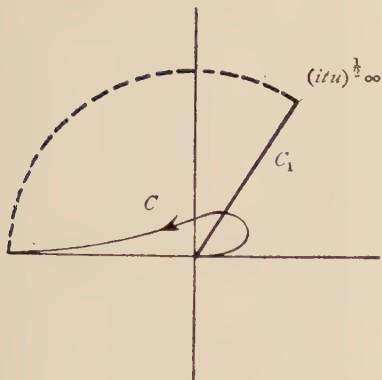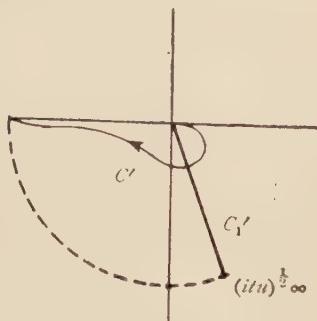
Fig. 1   Fig. 2

If $u < 0$, the path C_1' is transformed to the C' given by Fig. 2.

In the same way,

$$\varphi(u) = -\frac{u}{8} \int_{c-i\infty}^{c+i\infty} dt \frac{1}{t^2} \left(\frac{iu}{t}\right)^{1/2} H_1^{(2)} \left(\sqrt{\frac{iu}{t}}\right) \exp t .$$

Therefore (16) follows.

Equation (17) are obtained by expanding the Hankel function by its arguments and noticing the equations

$$\int_{c-i\infty}^{c+i\infty} dt \frac{1}{t^{n+1}} \exp t = \frac{2\pi i}{n!}$$

$$\int_{c-i\infty}^{c+i\infty} dt \frac{1}{t^{n+1}} \log t \cdot \exp t = \frac{2\pi i}{n!} \left(\sum_{S=1}^{n} \frac{1}{S} - \gamma\right).$$

Appendix C

$\varphi(u)$ satisfies the equation

$$(d^3/du^3 - k^3/u_2)\varphi = 0 \quad (k^3 = 1/4i).$$

The solutions at the large u are obtained easily

$$\varphi(u) = \exp(3ku^{1/3}) u^{2/3} \left(1 - \frac{8}{9} \frac{1}{ku^{1/3}} + \frac{20}{81} \frac{1}{(ku^{1/3})^2} + \cdots\right).$$

Three possible values for k are $(1/4)^{1/3}i$, $(1/4)^{1/3} \exp(-\pi/6 \cdot i)$ and $(1/4)^{1/3} \exp 7\pi i/6$. As is known from (13), $\varphi(u)/u$ tends to zero for the large u. So k must be taken to $(1/4)^{1/3}i$. From this (18) follows.

References

1) F. J. Dyson, Phys. Rev. **75** (1949), 1736.
2) for example
    Abdus Salam, Phys. Rev. **84** (1951), 426.
    J. C. Ward, Phys. Rev. **84** (1951), 894.
3) S. Sakata, H. Umezawa and S. Kamefuchi, Prog. Theor. Phys. **7** (1952), 377.
    H. Umezawa, Prog. Theor. Phys. **7** (1952), 551.
4) Ning Hu, Rhys. Rev. **80** (1950), 1109.
5) S. Kamefuchi and H. Umezawa, Prog. Theor. Phys. **9** (1953), 529.
6) S. F. Edwards, Phys. Rev. **90** (1953), 284.
7) Watson, *Theory of Bessel Function*, Cambridge (1922) 434.
8) R. J. Glauber, Phys. Rev. **84** (1951), 397.
9) G. C. Wick, Phys. Rev. **80** (1950), 268.
10) C. A. Hurst, Proc. Roy. Soc. **214** (1952), 44.
    R. Utiyama and T. Imamura, Prog. Theor. Phys. **9** (1953), 431.
11) F. J. Dyson, Phys. Rev. **73** (1948), 929.
12) J. M. Berger, L.L. Foldy and R. K. Osborne, Phys. Rev. **87** (1952), 1061.
13) J. Schwinger, Proc. N.A.S. **37** (1951), 452.

Studies on the Stochastic Problem of Electron-photon Cascades

Alladi RAMAKRISHNAN and P. M. MATHEWS

Department of Physics, University of Madras, Madras 25

A brief review of the well known fluctuation problem of cosmic radiation is given in a proper historical setting. We also present our latest numerical calculations for the second moments of the electron distribution as a sequel to those of Bhabha and Chakrabarthy for the mean number.

Introduction

The cascade theory* of cosmic ray showers first put forward by Bhabha and Heitler (A. 1) and independently by Carlson and Oppenheimer (A. 2) and subsequently developed by many authors, adequately describes the *average* behaviour of electron-photon cascades initiated by a fast electron or photon in its passage through matter. It has been recognised from the beginning that, since radiation and pair creation processes are governed by probability laws, cascade multiplication is essentially a stochastic process, a study of which can be complete only when considered from a statistical point of view. Since the Bethe-Heitler cross-sections for radiation and pair creation are continuous in the energy variable it has been realised that the main difficulty in such a treatment lies in the development of techniques to deal with the very difficult mathematical problem relating to the stochastic variable representing the number of particles distributed in a continuous infinity of states characterised by the energy parameter. This problem has attracted the attention of many workers and many powerful mathematical techniques have been devised. More than forty papers have been published but unfortunately in many of them, inadequate references have been made to earlier or contemporary work. Out of the large number of papers on this subject, only two deal in detail with numerical results relating to the electron-photon cascade—that of Scott and Uhlenbeck (B. 6) and of Janossy and Messel (D. 4). Thus the object of the present paper is twofold:

1. To give a brief summary of the developments of the mathematical techniques in their proper historical setting;

2. To present our latest numerical results** based upon the paper of Bhabha and Ramakrishnan (C. 4).

---

* At the end of this study, a comprehensive classified list of papers on this subject is given. In the course of the paper the references are indicated in brackets. These relate to the classified list.

** In a letter to this Journal (June 1953) we first announced some numerical results for the mean square number based on (C. 4). Since writing that letter, we discovered a mistake in (C. 4) which materially altered the results. The present calculations, made after correcting the mistake supersede the previous results.

Our results for the mean square number of electrons form a natural sequel to those for the mean number given by Bhabha and Chakrabarthy (A. 11). A brief comparison with the calculations of Scott and Uhlenbeck and those of Janossy and Messel will be made at the end of this paper.

PART I.

Mathematical techniques developed in the treatment of the fluctuation problem

Throughout this paper we ignore the lateral spread of the shower and treat the problem as one-dimensional as regards spatial distribution. The stochastic problem of cosmic ray showers can then be stated in the simplest mathematical terms as follows :

Statement of the problem

Given the initial energy spectrum of photons and electrons at thickness $t=0$ and that

(a) an electron of energy E radiates a quantum and has residual energy between E' and $E'+dE'$ with a probability per unit thickness $R^{(1)}(E', E)dE'$,

(b) a photon of energy E forms an electron pair one of which has energy between E' and $E'+dE'$ with probability per unit thickness $R^{(2)}(E', E)dE' = 2R'(E', E)dE'$, (The factor 2 occurs since we do not distinguish between positive and negative electrons.)

(c) an electron with energy E loses energy deterministically of magnitude $\beta(E)dt$ in its passage through a thickness dt.

We must calculate $\pi(n, E_1, m, E_2 ; t)$* the probability that there are n electrons with energy greater than E_1 and m photons with energy greater than E_2.

If we neglect process (c) i. e., ionisation, we shall call it Approximation A. If we include process (c) but assume β to be independent of E we shall it Approximation B. Cascade theories have been developed only under these approximations.

It is realised that $\pi(n, E_1, m, E_2 ; t)$ does not specify the statistical distribution in a manner satisfactory to the mathematician, but from a physical point of view it adequately describes the essential features of the process. The numerical evaluation of $\pi(n, E_1, m, E_2 ; t)$ itself is a large task which has yet to be achieved. The present numerical calculations are confined to the second moments of the number of electrons or photons above a specified energy.

The method that immediately suggests itself is the classical method which consists in expressing the 'state' of the system at $t+dt$ in terms of the 'state' at t or in other words in studying the behaviour of the system in an infinitesimal interval dt in terms of the 'state' of the system at t. This method is applicable only to Markovian processes.

---

* Throughout this paper the symbol π is used to denote a probability distribution function. When two or more functions are denoted by π, they will be distinguishable from the context,

But the function $\pi(n, E_1, m, E_2; t)$ does not define a Markoff process since the information that there are n electrons above E_1 and m photons above E_2 at t is insufficient to predict their behaviour in dt and a more complete description of the 'state' at t is necessary before we can apply the classical method. If the energy space were discrete, characterised by $E_1, E_2, \cdots, E_i, \cdots,$* then it would be possible to define a function

$$\pi(n_1, n_2, \cdots, n_i, \cdots, m_1, m_2, \cdots, m_i, \cdots; t)$$

representing the probability that there are n_i electrons and m_i photons in E_i, $i = 1, 2, \cdots$. This would be an adequate description of the 'state' of the system to predict its behaviour in interval dt. If the E-space is a continuum, no such function can be defined.

This difficulty was realised when the cascade theory was put forward, and so attempts were made to replace the cosmic ray process by simple statistical models. In their first paper, Bhabha and Heitler (A. 1) assumed a Poisson distribution for $N(E)$, the number of electrons above a specified energy E i. e., if $\pi(n, E; t)$ is the probability that there are n electrons above the energy E at t

$$\pi(n, E; t) = e^{-\lambda t}(\lambda t)^n / n!,$$
$$\varepsilon\{N^2(E)\} - [\varepsilon\{N(E)\}]^2 = \varepsilon\{N(E)\} = \lambda**. \qquad (1)$$

This was equivalent to the statement that the electrons in the various energy states are "uncorrelated"—an unreal assumption since the electrons are all generated from the initial parent, the incident particle. Furry (B. 1) was the first to take into account the multiplicative nature of the process and he obtained a distribution function on the basis of a simple multiplicative model which yielded fluctuations of much larger magnitude than the Poisson distribution. Assuming for simplicity, that there is only one type of particle having a probability B of splitting into two per unit distance of travel, he obtained an expression for the probability that there are n particles at t,

$$\pi(n; t) = \frac{1}{\varepsilon\{N\}}\left(1 - \frac{1}{\varepsilon\{N\}}\right)^{n-1}, \qquad (2)$$

where

$$\varepsilon\{N\} = e^{Bt},$$
$$\varepsilon\{N^2\} - [\varepsilon\{N\}]^2 = [\varepsilon\{N\}]^2 - \varepsilon\{N\}. \qquad (3)$$

Other attempts were made by Euler (B. 3), Nordseik, Lamb and Uhlenbeck (B. 5) and a lengthy discussion of this problem is to be found in the book of Arley (B.4).

The first attempt at a statistical treatment of the development of the shower taking

---

* There is no suggestion that $E_1, E_2, \cdots, E_n, \cdots$ are ordered. They merely indicate the discrete values of the energy in the entire range.

** Throughout this paper ε denotes expectation value.

into account the energy distribution of the particles was made by Scott and Uhlenbeck. To avoid the well known complications outlined earlier, they made the following assumptions.

(1)   Electrons and photons exist in discrete states each with a definite energy.

(2)   A limiting process exists by which the number of such states can be made as large as we please.

This in fact corresponds to the well known procedure in quantum mechanics in which one considers all the phenomena to take place in a large box of dimesion L. The states of the electrons and photons are then discrete and their number tends to infinity as L tends to infinity. Scott and Uhlenbeck defined a 'Master function' $\pi(n_1, n_2, \cdots, m_1, m_2, \cdots; t)$ representing the probability there are n_i electrons and m_i photons with energy $E_i, i = 1, 2, \cdots$ and from these other statistical functions are defined by taking suitable average values. An identical procedure was adopted by Bhabha and Ramakrishnan* which ultimately led them independently to the formulation of the " theory of product densities " which deals directly with the continuous case and yields the same equations without the aid of a laborious limiting process. The equations of Bhabha and Ramakrishnan differ from those of Scott and Uhlenbeck only in that they take into account the Bethe-Heitler cross sections but their paper goes further in strictly carrying the calculations to the end.

Besides the theory of product densities, three other powerful techniques have been developed in the treatment of this problem which we shall consider in order of their historical development.

(1) *The method of product density functions*

This method, independently put forward by Bhabha (C. 2) and Ramakrishnan (C. 1) can be briefly stated as follows. Here we use the notation of Ramakrishnan.

If particles are distributed in the E-space which is a continuum, we define a function $f_n(E_1, E_2, \cdots, E_n)$ called the product density of degree n, such that $f_n(E_1, E_2, \cdots, E_n)$ $\cdot dE_1 dE_2 \cdots dE_n$ represents the probability that there is a particle in dE_1, a particle in $dE_2, \cdots$, and a particle in dE_n. f_n is not a probability density but $f_n(E_1, E_2, \cdots, E_n)$ $\cdot dE_1 dE_2 \cdots dE_n$ is a probability magnitude.

If $N(E)$ is the stochastic variable representing the number of particles with energy above E then

$$\varepsilon \{ [N(E) - N(E_0)]^r \} = \sum_{s=1}^{r} C_s^r \int_E^{E_0} \int_E^{E_0} \cdots \int_E^{E_0} f_s(E_1, E_2, \cdots, E_s) dE_1 dE_2 \cdots dE_s , \quad (4)$$

where C_s^r are coefficients defined by the identity

$$M^r = \sum_{s=1}^{r} C_s^r M(M-1) \cdots (M-s+1) , \qquad (5)$$

M being a positive integer. The coefficients are obtained by putting $M = 1, 2, \cdots$.

---

*   For a full account of this, see Alladi Ramakrishnan, Ph. D. Thesis (1951) University of Manchester.

It is clear from (4) that for the calculation of the n-th moment of $[N(E) - N(E_0)]$, we should know the product densities of degree less than or equal to n. In the cosmic ray problem since we are dealing with two types of particles, electrons and photons, we define $f_n(E_1, E_2, \cdots, E_n; t)$ and $g_n(E_1, E_2, \cdots, E_n; t)$ as the product densities of degree n of electrons and photons respectively and $fg_{n,m}(E_1, E_2, \cdots, E_n; E_{n+1}, \cdots, E_{n+m}; t)$ as the mixed product density of degree (n, m) of electrons and photons, i. e., $fg_{n,m}(E_1, E_2, \cdots, E_n; E_{n+1}, \cdots, E_{n+m}; t) dE_1 dE_2 \cdots dE_{n+m}$ is the probability that there is an electron in dE_1, an electron in dE_2, $\cdots$, and an electron in dE_n, a photon in dE_{n+1}, a photon in dE_{n+2}, $\cdots$, and a photon in dE_{n+m}.

Then it can be shown that since $f_n(E_1, E_2, \cdots, E_n; t) dE_1 dE_2 \cdots dE_n$ is a probability magnitude, its variation with t can be expressed in terms of the f_i, g_i, and $fg_{i,j}$, funtions. Bhabha has formally written down the equations for f_n, g_n, and $fg_{n,m}$ but equations involving product densities of order greater than two are not of immediate interest since the analytical solution of f_2, g_2 and $fg_{1,1}$, is itself a long and tedious process which has been obtained by Bhabha and Ramakrishnan under approximation A. Section II of the paper deals with numerical results based upon their solution.

We must admit that even if we know f_n, g_n, and $fg_{n,m}$ for every n, m, our knowledge of the statistical distribution of electrons or photons is still incomplete. A more complete description is achieved by the means of the characteristic functional which will be discussed a little later.

(2) *Method of regeneration points*

To get round the difficulties now outlined, Janossy (D. 2) employed an entirely different technique which, it was realised later, was employed previously by Bellman and Harris (D. 1) and C. Palm (D. 2) in relation to simpler stochastic processes. A complete account of the historical evolution of this " method of regeneration points " is given in a paper of Bartlett and Kendall (D. 5). But it is to Janossy we owe its ingenious application to the vexed problem of cosmic radiation.

Janossy[*] defined a function $\Phi^{(i)}(E, N; t; E_0)$ as the probability that at thickness t there are N electrons each with energy greater than E where E_0 is the initial energy of the incident particle. $i = 1$ refers to an electron-initiated shower while $i = 2$ refers to a photon-initiated shower. The method of regeneration points uses the following argument : somewhere at a thickness less than t (and this is called a regeneration point) the incident particle may radiate a pair and these secondaries and primary together become indpendent primaries of stochastic processes. By this argument we are led to the following integral equation

$$\Phi^{(i)}(E, N; t; E_0) = \sum_{N'+N''=N} \int_0^t d\tau \int_E^{E_0} e^{-a_i(E_0)\tau} R^{(i)}(E', E_0) \Phi^{(1)}(E, N'; t-\tau; E').$$

$$\Phi^{(3-i)}(E, N''; \ t-\tau; \ E_0-E')dE' + \Delta(N-2+i)e^{-a_i(E_0)t} \ , \qquad (6)$$

where m a non-negative integer; $\Delta(m)=1$ if $m=0$, and otherwise zero.

At $t=0$*

$$\Phi^{(1)}(E, 1; \ t; \ E_0)=1 \ ,$$

$$\Phi^{(1)}(E, N; \ t; \ E_0)=0 \quad \text{if} \quad N \neq 1 \ ,$$

$$\Phi^{(2)}(E, 0; \ t; \ E_0)=1 \ ,$$

$$\Phi^{(2)}(E, N; \ t; \ E_0)=0 \quad \text{if} \quad N \neq 0 \ . \qquad (7)$$

Differentiating the above equation with respect to t, we have

$$\frac{\partial \Phi^{(i)}(E, N; \ t; \ E_0)}{\partial t} = - \Phi^{(i)}(E, N; \ t; \ E_0)a_i(E_0)$$

$$+ \sum_{N'+N''=N} \int_E^{E_0} \Phi^{(i)}(E, N'; \ t; \ E') \Phi^{(3-i)}(E, N''; \ t; \ E_0-E') R^{(i)}(E', E_0)dE'. \qquad (8)$$

It was noticed by Ramakrishnan** that for processes homogeneous and Markovian with respect to t, the above method consists in considering the change in the first infinitesimal interval 0 to $\triangle$, and expressing the state of the system at t in terms of the state at $\triangle$ as contrasted with the classical method of expressing the state at $t+\triangle$ in terms of the state at t. The main advantage of this method is that it is simpler to study the change in the system during the first infinitesimal interval $\triangle$. The striking contrast between these two approaches was observed much earlier by Kolmogoroff*** when he derived the backward and forward differential equations of the function $\pi(s_j; \ t; \ s_i)$ representing the probability that a system is in the state s_j at t given that it was in s_i at $t=0$. In any stochastic process if we interpret the 'state' of the system suitably, the backward and forward differential equations may be written down directly. $\pi(s_j; \ t; \ s_i)$ is defined by the Kolmogoroff equations only if we know $R(s_j, s_i)$ where $\pi(s_j; \ t; \ s_i) \to R(s_j, s_i)t$ as $t \to 0 (j \mid i)$. But there may occur stochastic processes *where it is not possible to know the limit $R(s_j, s_i^e)$ for all but a few simple s_i's.*

At $t=0$, the state of the system s_i will be defined in a simpler manner than at some t, and so it may be possible to study the transitions from this intial state but not those from the state at t. In such a case we can write only the backward differential equations of the process.

It is to be noted that even for product densities backward differential equations can be obtained. But in this case, it is easier to solve the forward equations of Bhabha and Ramakrishnan since the initial conditions of the product densities is the same inside and outside the integral sign in their integro-differential equations.

---

* In reference (D. 6) the initial conditions stated are wrong, but fortunately the mistake was not carried over into the subsequent calculations.

** See M. S. Bartlett, reference (D. 4).

*** For an account of this method see reference (H. 4).

(3)   *The method of Janossy functions*

In an attempt to treat the fluctuation problem of nucleon cascades* by a method different from that outlined just now, Janossy (E. 1) defined a function $\psi_h^3(E_1, E_2, \cdots, E_h)$ $\cdot dE_1 dE_2 \cdots dE_h$ representing the probability that there is one particle in dE_1, one in dE_2, $\cdots$, and one in dE_h and *none in the* other energy states.   Note that ψ_h is not a probability density.   Janossy in fact dealt with $\psi_h/h!$ but it is ψ_h that has a physical significance since the particles are indistinguishable.

The Janossy functions ψ_h bear a close relationship to the product density function** f_h and we give here without proof the following result

$$f_h(E_1, E_2, \cdots, E_h) = \sum_h \frac{1}{(n-h)!} \int_{E_n} \int_{E_{n-1}} \cdots \int_{E_{h+1}} \psi_n(E_1, E_2, \cdots, E_n) dE_n dE_{n-1} \cdots dE_{h+1}.$$

$$(9)$$

Forward differential equations for ψ_h can be written down directly but equations for ψ_h will involve $\psi_m(m > h)$.   In the case of product densities equations for f_h will involve densities of order less than h and hence successive solution of f_h is possible.

The equations of the nucleon cascade problem have been derived using the two methods and their equivalence established by Ramakrishnan (G. 1) and by Messel and Potts (G. 2).

In the case of the electron-photon cascade, it is clear that we have to define $\psi_{n,m}(E_1, E_2, \cdots, E_n; E_{n+1}, E_{n+2}, \cdots, E_{n+m}) dE_1 \cdots dE_{n+m}$ as the probability that there is an electron in dE_1, one in $dE_2, \cdots$, one in dE_n, a photon in dE_{n+1}, a photon in $dE_{n+2}, \cdots$, and a photon in dE_{n+m}.   This function was used extensively by Messel and his collaborators.

(4)   *The characteristic functional method*

The characteristic functional is a powerful mathematical device to study the joint distribution of a continuous infinity of stochastic variables.   If the stochastic variable $N(E)$ represents the number of particles with energy above E then symbolically $-dN(E)$ is a stochastic variable which represents the number in dE.   Then, if

$$C[\theta(E), t] = \varepsilon \{ e^{-i\int_E \theta(E) dN(E, t)} \},$$

$$(10)$$

$C[\theta(E), t]$ is called the characteristic functional and $\theta(E)$ its argument.   " *The characteristic functional, when it can be determined contains in portmanteau form all the probability relationships associated with the state of the ' population ' of particles at t.*"

The characteristic functional in the cosmic ray problem must necessarily involve two argument functions and so we shall write

---

* The problem of nucleon cascades differs from that of electron-photon cascades in that we deal with only one type of particle.

** For proof of (9), see reference (G. 1).

$$C^{(i)}[\theta(E), \chi(E) ; t] = \varepsilon \{ e^{-i\int_E [\theta(E)dN^{(i)}(E,t) + \chi(E)dM^{(i)}(E,t)]} \}. \qquad (11)$$

$-dN^{(i)}(E,t)$ and $-dM^{(i)}(E,t)$ are stochastic variables representing respectively the number of electrons and photons in dE. $i=1$ refers to an electron-initiated shower, $i=2$ to a photon-initiated shower.

By using the now familiar arguments of the regeneration point method, Bartlett and Kendall (D. 5) obtained the equation

$$C^{(i)}[\theta(E), \chi(E) ; t; E_0] = \int_0^{E_0} \int_0^t e^{-a_i(E_0)\tau} C^{(i)}[\theta(E), \chi(E) ; t-\tau ; E'],$$

$$C^{(3-i)}[\theta(E), \chi(E) ; t-\tau ; E_0 - E'] R^{(i)}(E', E_0) dE' d\tau + e^{-a_i(E_0)t} e^{i\theta(E_0)}.$$

$$(12)$$

Putting

$$\chi(E) = 0$$

and

$$\theta(E') \equiv 0, \ (E' < E)$$

$$\equiv -i \log u, \ (E' > E) \qquad (13)$$

we get Janossy's equation (6).

Differentiating with respect to t we get

$$\frac{\partial C^{(i)}[\theta(E), \chi(E) ; t; E_0]}{\partial t} = -C^{(i)}[\theta(E), \chi(E) ; t; E_0] a^{(i)}(E_0)$$

$$+ \int_0^{E_0} C^{(i)}[\theta(E), \chi(E) ; t; E'] C^{(3-i)}[\theta(E), \chi(E) ; t; E_0 - E'] R^{(i)}(E', E_0) dE'.$$

$$(14)$$

(5) *Analytical solutions of the cascade equations*

The analytical solutions of the equations by various methods have been attempted only in the case when $R^{(i)}(E', E)dE'$ can be expressed in the form $R^{(i)}(q)dq$, $q = E'/E$. The method of attack essentially consists in using the Mellin's transform technique and reducing the integro-differential equations to differential equations. We here summarise the various attempts that have been made.

Bhabha and Ramakrishnan have obtained explicit analytical solutions for product-densities of degree two. The solutions of equations involving product densities of higher order have yet to be obtained.

Janossy attempted to solve his equation by using the method of generating functions but has succeeded only in obtaining formal expressions for the moments of the distribution. Messel and Gardner (F. 12) have claimed that the formal solution of Janossy's equation can be obtained without the aid of generating functions.

Messel and his co-workers have made a series of contributions to the analytical solution of the cascade equations (including ionisation loss) based on Janossy's functions by an ex-

tensive use of matrix methods.

The solution of the characteristic functional equation is yet to be attempted.

PART II.

The analytical solution of Bhabha and Ramakrishnan

As has been pointed out in the introduction, while the formal mathematical treatment of the fluctuation problem has been completed, numerical calculations have been confined only to the second moments of the electron distribution in the case of *Approximation A**. We present here our latest numerical results based on the product density equations of Bhabha and Ramakrishnan (C. 4). We shall summarise their results to the extent necessary for a proper appreciation of our numerical calculations. For the sake of completeness we will reproduce the product density equtions.

If f_1 and f_2 are the product densities of electrons of order one ond two, g_1 and g_2 of photons of order one and two, and $fg_{1,1}$ the mixed product density of electrons and photons, they satisfy the following equtions**.

Product-density-Degree 1 : —

$$\frac{\partial f_1(E; t)}{\partial t} = - f_1(E; t)\int_0^E R^{(1)}(E', E)dE' + \int_E^\infty f_1(E'; t)R^{(1)}(E, E')dE'$$

$$+ 2\int_E^\infty g_1(E'; t)R^{(2)}(E, E')dE', \qquad (15)$$

$$\frac{\partial g_1(E; t)}{\partial t} = -g_1(E; t)\int_0^E R^{(2)}(E', E)dE' + \int_E^\infty f_1(E'; t)R^{(1)}(E'-E, E')dE', \qquad (16)$$

As stated earlier, $R^{(2)}(E', E) = 2R'(E', E)$.

Product density-Degree 2 : —

$$\frac{\partial f_2(E_1, E_2; t)}{\partial t} = -f_2(E_1, E_2; t)\left\{\int_0^{E_1} R^{(1)}(E, E_1)dE + \int_0^{E_2} R^{(1)}(E, E_2)dE\right\}$$

$$+ \int_{E_1}^\infty f_2(E, E_2; t)R^{(1)}(E_1, E)dE + \int_{E_2}^\infty f_2(E, E_1; t)R^{(1)}(E_2, E)dE$$

$$+ 2\int_{E_2}^\infty fg_{1,1}(E_1, E; t)R^{(2)}(E_2, E)dE + 2\int_{E_1}^\infty fg_{1,1}(E_2, E; t)R^{(2)}(E_1, E)dE$$

$$+ 2g_1(E_1+E_2; t)R^{(2)}(E_1, E_1+E_2), \qquad (17)$$

---

*   Messel and Gardner (F. 12) have recently announced that they are planning a programme of numerical work on the Ferranti electronic computer at the University of Toronto computation centre.

**   For the derivation of these equations see (C. 1). It is to be noted that we have used a notation for the cross-sections different from that of (C. 1) and also of (A. 11).

$$\frac{\partial fg_{1,1}(E_1, E_2;\ t)}{\partial t} = -fg_{1,1}(E_1, E_2;\ t)\left\{\int_0^{E_1} R^{-1}(E, E_1)dE + \int_0^{E_2} R^{2}(E, E_2)dE\right\}$$

$$+ \int_{E_1}^{\infty} fg_{1,1}(E, E_2;\ t)R^{(1)}(E_1, E)dE + \int_{E_1}^{\infty} 2g_2(E, E_2;\ t)R^{(2)}(E_1, E)dE$$

$$+ \int_{E_2}^{\infty} f_2(E_1, E;\ t)R^{(1)}(E-E_2, E)dE + f_1(E_1+E_2;\ t)R^{(1)}(E_1, E_1+E_2), \qquad (18)$$

$$\frac{\partial fg_{1,1}(E_2, E_1;\ t)}{\partial t} = -fg_{1,1}(E_2, E_1;\ t)\left\{\int_0^{E_2} R^{-1}(E, E_2)dE + \int_0^{E_1} R^{(2)}(E, E_1)dE\right\}$$

$$+ \int_{E_2}^{\infty} fg_{1,1}(E, E_1;\ t)R^{(1)}(E_2, E)dE + \int_{E_2}^{\infty} 2g_2(E, E_1;\ t)R'(E_2, E)dE$$

$$+ \int_{E_1}^{\infty} f_2(E_2, E;\ t)R^{-1}(E-E_1, E)dE + f_1(E_1+E_2;t)R^{-1}(E_2, E_1+E_2), \qquad (19)$$

$$\frac{\partial g_2(E_1, E_2;\ t)}{\partial t} = -g_2(E_1, E_2;\ t)\left\{\int_0^{E_1} R^{(2)}(E, E_1)dE + \int_0^{E_2} R^{(2)}(E, E_2)dE\right\}$$

$$+ \int_{E_2}^{\infty} fg_{1,1}(E, E_1;\ t)R^{(1)}(E-E_2, E)dE + \int_{E_1}^{\infty} fg_{1,1}(E, E_2;\ t)R^{-1}(E-E_1, E)dE.$$
$$\tag{20}$$

The initial conditions of the problem are defined thus :
At $t=0$, there is only one primary of energy E_0.   Therefore at $t=0$

$$f_1(E;\ 0) = \delta(E-E_0),\ g_1(E;\ 0) = 0,$$

$$f_2(E_1, E_2;\ 0) = 0,\ g_2(E_1, E_2;\ 0) = 0,\ fg_{1,1}(E_1, E_2;\ 0) = 0,$$

$R^{(1)}(E'\ E)$ and $R'(E', E)$ are the Bethe-Heitler cross-sections given by

$$R^{(1)}(E', E) = \left\{\frac{E-E'}{E} - \left(\frac{4}{3}+a\right)\left(1 - \frac{E}{E-E'}\right)\right\}\frac{1}{E},$$

$$R'(E', E) = \left\{1 - \left(\frac{4}{3}+a\right)\left(\frac{E'}{E} - \frac{E'^2}{E^2}\right)\right\}\frac{1}{E}. \qquad (22)$$

(15) and (16) are the well known equations of the cascade theory for the mean number.
(17) to (20) have been completely solved by Bhabha and Ramakrishnan using the Mellin's
transforms,

$$\nu(r;\ t) = \int_0^{\infty} E^{r-1}f_1(E;\ t)dE,\quad \gamma(r;\ t) = \int_0^{\infty} E^{r-1}g_1(E;\ t)dE,$$

$$\nu_2(r, s;\ t) = \int_0^{\infty} dE \int_0^{\infty} dE'\cdot E^{r-1}E'^{s-1}f_2(E, E';\ t),$$

$$\nu_r(r, s;\ t) = \int_0^{\infty} dE \int_0^{\infty} dE'\cdot E^{r-1}E'^{s-1}fg_{1,1}(E, E';\ t),$$

$$\gamma_2(r, s;\ t) = \int_0^{\infty} dE \int_0^{\infty} dE'\cdot E^{r-1}E'^{s-1}g_2(E, E';\ t). \qquad (23)$$

They obtained the following equations* (written in matrix form)

$$\frac{d}{dt}\phi(r;t)=\tau(r)\phi(r;t),\qquad(24)$$

where

$$\phi(r;t)=\begin{pmatrix}\nu(r;t)\\\gamma(r;t)\end{pmatrix}$$

and

$$\tau(r)=\begin{pmatrix}-A_r & B_r\\C_r & -D\end{pmatrix},\qquad(25)$$

and

$$A_r=\left(\frac{4}{3}+\alpha\right)\left\{\frac{d}{dr}\log\Gamma(r)+\gamma-1+\frac{1}{r}\right\}+\frac{1}{2}-\frac{1}{r(r+1)},$$

$$B_r=2\left\{\frac{1}{r}-\left(\frac{4}{3}+\alpha\right)\frac{1}{(r+1)(r+2)}\right\},$$

$$C_r=\frac{1}{r+1}+\left(\frac{4}{3}+\alpha\right)\frac{1}{r(r-1)},$$

$$D=\frac{7}{9}-\frac{1}{6}\alpha.\qquad(26)$$

α is a constant characteristic of the material** on which the initial electron is incident. γ inside (26) represents the Euler-Mascheroni constant. The matrix $\tau(r)$ has two eigen values $-\lambda_r$ and $-\mu_r$ given by

$$\begin{pmatrix}\lambda_r\\\mu_r\end{pmatrix}=\frac{1}{2}(A_r+D)\mp\frac{1}{2}\{(A_r-D)^2+4B_rC_r\}^{1/2}.\qquad(27)$$

If

$$\phi^{(2)}(r,s;t)=\begin{pmatrix}\nu_2(r,s;t)\\\nu\gamma(r,s;t)\\\nu\gamma(s,r;t)\\\gamma_2(r,s;t)\end{pmatrix},\qquad(28)$$

then

$$\frac{d}{dt}\phi^{(2)}(r,s;t)=\{\tau(r)\times1+1\times\tau(s)\}\phi^{(2)}(r,s;t)+\psi(r,s;t),\qquad(29)$$

where 1 is the 2×2 unit matrix and ψ is a column vector defined by

---

* As regards the transform equations, the notation is identical with that used in reference (C. 4). The paper of Bhabha and Ramakrishnan contains a number of mistakes which have all been pointed out in the Appendix.

** In this paper we take the value of α to be .0246.

$$\psi(r, s;\ t) = \psi_1(r, s)\, E_0^{\,r+s-2}\, e^{-\lambda_{r+s-1}t} + \psi_2(r, s)\, E_0^{\,r+s-2}\, e^{-\mu_{r+s-1}t}, \qquad (30)$$

where

$$\psi_1(r, s) = \frac{1}{\mu_{r+s-1} - \lambda_{r+s-1}}\begin{vmatrix} a_1(r, s)\, C_{r+s-1} \\ a_2(r, s)\, (D - \lambda_{r+s-1}) \\ a_2(s, r)\, (D - \lambda_{r+s-1}) \\ 0 \end{vmatrix},$$

$$\psi_2(r, s) = \frac{1}{\mu_{r+s-1} - \lambda_{r+s-1}}\begin{vmatrix} -a_1(r, s)\, C_{r+s-1} \\ a_2(r, s)\, (\mu_{r+s-1} - D) \\ a_2(s, r)\, (\mu_{r+s-1} - D) \\ 0 \end{vmatrix}, \qquad (31)$$

and

$$a_1(r, s) = 2\frac{\Gamma(r)\,\Gamma(s)}{\Gamma(r+s)}\left\{ 1 - \left(\frac{4}{3} + a\right)\frac{rs}{(r+s+1)(r+s)}\right\},$$

$$a_2(r, s) = \frac{\Gamma(r)\,\Gamma(s)}{\Gamma(r+s)}\left\{ \frac{s}{r+s} + \left(\frac{4}{3} + a\right)\frac{r}{s-1}\right\}. \qquad (32)$$

The initial conditions of the problem require that

$$\nu(s;\ 0) = E_0^{\,s-1}, \quad \gamma(s;\ 0) = 0. \qquad (32)$$

Thus solving (24) we obtain

$$f_1(E;\ t) = \frac{1}{2\pi i E_0}\int_{\sigma-i\infty}^{\sigma+i\infty}\left(\frac{E_0}{E}\right)^s\left\{ \frac{D - \lambda_s}{\mu_s - \lambda_s}\, e^{-\lambda_s t} + \frac{\mu_s - D}{\mu_s - \lambda_s}\, e^{-\mu_s t}\right\}ds. \qquad (34)$$

The mean number of electrons with energy greater than E is given by

$$\varepsilon\{N(E)\} = \frac{1}{2\pi i}\int_{\sigma-i\infty}^{\sigma+i\infty}\left(\frac{E_0}{E}\right)^{s-1}\frac{1}{s-1}\left\{ \frac{D - \lambda_s}{\mu_s - \lambda_s}\, e^{-\lambda_s t} + \frac{\mu_s - D}{\mu_s - \lambda_s}\, e^{-\mu_s t}\right\}ds. \qquad (35)$$

For all but small thickness we neglect the term involving $e^{-\mu_s t}$.

The solution of the inhomogeneous equation (29) is given by

$$\phi^{(2)}(r, s;\ t) = \zeta(r, s;\ t)\int_0^t \zeta^{-1}(r, s;\ t)\psi(r, s;\ t)dt. \qquad (36)$$

when ζ is the solution of the homogeneous equation

$$\frac{d}{dt}\zeta(r, s;\ t) = \{\tau(r)\times 1 + 1\times \tau(s)\}\zeta(r, s;\ t). \qquad (37)$$

Corresponding to the two parts ψ_1 and ψ_2 of ψ, $\phi^{(2)}$ itself splits into two parts $\phi_1^{(2)}$ and $\phi_2^{(2)}$. Since the only t dependent factors of the integral in (36) are the exponentials, we find easily

$$\phi_1^{(2)}(r, s;\ t) = E_0^{\,r+s-2}\zeta_0(r, s)\eta(\lambda_{r+s-1})\zeta_0^{-1}(r, s)\psi_1(r, s), \qquad (38)$$

$$\phi_2^{(2)}(r, s; t) = E_0^{r+s-2}\zeta_0(r, s)\eta(\mu_{r+s-1})\zeta_0^{-1}(r, s)\psi_2(r, s), \tag{39}$$

where $\eta(x)$ is the diagonal matrix with the four elements

$$\frac{e^{-xt} - e^{-(\lambda_r + \lambda_s)t}}{\lambda_r + \lambda_s - x}, \quad \frac{e^{-xt} - e^{-(\lambda_r + \mu_s)t}}{\lambda_r + \mu_s - x}, \quad \frac{e^{-xt} - e^{-(\mu_r + \lambda_s)t}}{\mu_r + \lambda_s - x}, \quad \frac{e^{-xt} - e^{-(\mu_r + \mu_s)t}}{\mu_r + \mu_s - x}. \tag{40}$$

We note a very important feature of this matrix. *It remains finite even if any of the denominators of its elements tends to zero.* This has to be taken into account in tabulating $\phi^{(2)}$ for numerical calculations using the saddle point method.

From the theory of product densities the mean square number of electrons with energy greater than E is given by

$$\varepsilon\{N^2(E; t)\} = \varepsilon\{N(E; t)\} + \int_E^{E_0} \int_E^{E_0} f_2(E_1, E_2; t) dE_1 dE_2, \tag{41}$$

$$f_2(E_1, E_2; t) = \frac{1}{(2\pi i E_0)^2} \int_{\sigma-i\infty}^{\sigma+i\infty} \int_{\sigma-i\infty}^{\sigma+i\infty} \nu_2(r, s; t) E_1^{-r} E_2^{-s} dr\, ds. \tag{42}$$

Bhabha and Ramakrishnan have calculated this (see equation (41 a) of their paper) and we here give the suitable approximation for all but small thicknesses. This approximation results from the fact that terms involving $e^{-(\lambda_r + \lambda_s)t}$ and $e^{-\lambda_{r+s-1}t}$ exceed the others (i. e. those involving μ in the exponent i. e. $e^{-(\mu_r + \mu_s)t}$, $e^{-\mu_{r+s-1}t}$, $e^{-(\lambda_r + \mu_s)t}$, $e^{-(\lambda_s + \mu_r)t}$) by an order of magnitude for all but small t. However, it may seem that even for small thicknesses the terms involving μ in the exponent have to be taken into account when the denominators of the matrix occurring in (40) vanish. Fortunately in the evaluation of (43) by the double saddle-point method, we find that only $\lambda_r + \lambda_s - \lambda_{r+s-1}$ vanishes at some point in the domain of the saddle-point corresponding to the physically important values of E and so terms involving μ in the exponent can be neglected.

$$\varepsilon\{N^2(E; t)\} = \varepsilon\{N(E; t)\} + \frac{1}{(2\pi i)^2} \int_{\sigma-i\infty}^{\sigma+i\infty} \int_{\sigma-i\infty}^{\sigma+i\infty} \left(\frac{E_0}{E}\right)^{r+s-2} \frac{G(r, s; t)e^{-(\lambda_r + \lambda_s)t}}{(r-1)(s-1)} dr\, ds, \tag{43}$$

where

$$G(r, s; t) = \frac{G_1(r, s) + G_2(r, s)e^{(\lambda_r + \lambda_s - \lambda_{r+s-1})t}}{\lambda_r + \lambda_s - \lambda_{r+s-1}}, \tag{44}$$

and*

$$G_1(r, s) = \frac{D - \lambda_r}{\mu_r - \lambda_r} \cdot \frac{D - \lambda_s}{\mu_s - \lambda_s} \cdot \frac{1}{\lambda_r + \lambda_s - \mu_{r+s-1}} \left\{ a_1(r, s) C_{r+s-1} \right.$$

$$\left. + a_2(r, s)\frac{\mu_s - D}{C_s}(D - \lambda_r - \lambda_s) + a_2(s, r)\frac{\mu_r - D}{C_r}(D - \lambda_r - \lambda_s) \right\}, \tag{45}$$

---

* In their paper Bhabha and Ramakrishnan omitted by mistake the factor $[(\mu_s - \lambda_s)(\mu_r - \lambda_r)]^{-1}$ occurring in (45). Their expression (41) must be multiplied by that factor.

$$G_2(r, s) = \Big[\{D(D-x)(x_r+x_s) + \lambda_r \mu_r x_s + \lambda_s \mu_s x_r - x x_r x_s\} a_1(r, s) C_{r+s-1}$$

$$+ \frac{(D-\lambda_s)(\mu_s-D)}{C_s} \{D(x_r+x_s) - \lambda_r \mu_r + (\lambda_s-x)(\mu_s-x)\} a_2(r, s)(D-\lambda_{r+s-1})$$

$$+ \frac{(D-\lambda_r)(\mu_r-D)}{C_r} \{D(x_r+x_s) - \lambda_s \mu_s + (\lambda_r-x)(\mu_r-x)\} a_2(s, r)(D-\lambda_{r+s-1}) \Big].$$

$$\{(\mu_{r+s-1}-\lambda_{r+s-1})(\lambda_r+\mu_s-x)(\lambda_s+\mu_r-x)(\mu_r+\mu_s-x)\}^{-1}, \tag{46}$$

with $x=\lambda_{r+s-1}$ and $x_r=\lambda_r+\mu_r-x$.

We have expressed $G(r, s; t)$ in the form (44) to facilitate tabulation of G. The factor $\dfrac{1}{\lambda_r+\lambda_s-\lambda_{r+s-1}}$ has been separated out since it becomes infinite at points where $\lambda_r+\lambda_s=\lambda_{r+s-1}$. At these points, as has already been mentioned, $G_1+G_2 e^{(\lambda_r+\lambda_s-\lambda_{r+s-1})t}$ tends to zero in such a manner that G is finite.

Numerical evaluation of $\varepsilon\{N^2(E; t)\}$

$\varepsilon\{N(E; t)\}$ and $\varepsilon\{N^2(E; t)\}$ are functions of E/E_0 if E_0 is the energy of the incident electron. We then write them as $\varepsilon\{N(y; t)\}$ and $\varepsilon\{N^2(y; t)\}$ respectively where $y=\log(E_0/E)$. Thus we have

$$\varepsilon\{N^2(y; t)\} = \varepsilon\{N(y; t)\} + I(y; t), \tag{47}$$

where

$$I(y; t) = \frac{1}{(2\pi i)^2} \int_{\sigma-i\infty}^{\sigma+i\infty}\int_{\sigma-i\infty}^{\sigma+i\infty} \frac{e^{y(r+s-2)}}{(r-1)(s-1)} G(r, s; t)e^{-(\lambda_r+\lambda_s)t} dr\, ds. \tag{48}$$

To apply the saddle-point method for the numerical evaluation of the complex integral, we write $I(y; t)$ in the form

$$I(y; t) = \frac{1}{(2\pi i)^2} \int_{\sigma-i\infty}^{\sigma+i\infty}\int_{\sigma-i\infty}^{\sigma+i\infty} e^{w(r,s;t)} dr\, ds, \tag{49}$$

where

$$w(r, s; t) = y(r+s-2) - t(\lambda_r+\lambda_s) + \log G(r, s; t)$$

$$-\log(r-1) - \log(s-1). \tag{50}$$

The saddle point (r_0, s_0) is then determined by the equations

$$\frac{\partial w(r, s; t)}{\partial r} = \frac{\partial}{\partial r}\log G(r, s; t) - \frac{1}{r-1} - t\frac{d\lambda_r}{dr} + y = 0, \tag{51}$$

$$\frac{\partial w(r, s; t)}{\partial s} = \frac{\partial}{\partial s}\log G(r, s; t) - \frac{1}{s-1} - t\frac{d\lambda_s}{ds} + y = 0. \tag{52}$$

G is symmetric in r and s. So also is w. Hence $r_0=s_0$ and the saddle points for variations of y and t lie on $r=s$.

The first step was to find the saddle-points of $e^{\omega(r,s;t)}$ for different values of y for each t. For finding this point, the differential co-efficients $\dfrac{\partial}{\partial r}(\log G)$ were evaluated at $r=s$. A table of $t\dfrac{d\lambda_r}{dr}+\dfrac{1}{r-1}-\dfrac{\partial}{\partial r}(\log G)$ against r, $(r=s)$ was formed and inverse interpolation carried out to determine the saddle points for various values of y and t. The choice of y is restricted by the fact that the saddle point should fall in the region of r, s in which tabulation is done. The range depends on t.

The saddle points having been determined thus, it was necessary to find $\dfrac{\partial^2 \omega}{\partial r^2}$ and $\dfrac{\partial^2 \omega}{\partial r \partial s}$ at these points.

We notice

$$\frac{\partial^2 \omega(r,s;t)}{\partial r^2}=\frac{\partial^2 \omega(r,s;t)}{\partial s^2} \quad \text{at} \quad r=s, \tag{53}$$

$$\left[\frac{\partial^2 \omega(r,s;t)}{\partial r^2}\right]_{r=s}=\left[\frac{\partial^2 \log G(r,s;t)}{\partial r^2}\right]_{r=s}-t\,\frac{d\lambda_r^2}{dr^2}+\frac{1}{(r-1)^2}. \tag{54}$$

To calculate $\dfrac{\partial^2 \omega}{\partial r \partial s}$ it is not necessary to tabulate $\dfrac{\partial \omega}{\partial r}$ at all r, s if we use the symmetry property of $\omega(r,s;t)$.

If we define

$$\Omega(r;t)=\omega(r,r;t)=\log G(r,r;t)-2\lambda_r t+2y(r-1)-2\log(r-1), \tag{55}$$

$$\frac{\partial\Omega(r;t)}{\partial r}=2\left[\frac{\partial\omega(r,s;t)}{\partial r}\right]_{r=s}, \tag{56}$$

$$\left[\frac{\partial^2 \omega(r,s;t)}{\partial r \partial s}\right]_{r=s}=\frac{1}{2}\frac{\partial^2 \Omega(r;t)}{\partial r^2}-\left[\frac{\partial^2 \omega(r,s;t)}{\partial r^2}\right]_{r=s}$$

$$=\frac{1}{2}\frac{\partial^2 \log G(r,r;t)}{\partial r^2}-\left[\frac{\partial^2 \log G(r,s;t)}{\partial r^2}\right]_{r=s}, \tag{57}$$

$\dfrac{\partial^2}{\partial r^2}(\log G(r,r;t))$ can be evaluated directly by using the standard formula for obtaining the differential co-efficients of a tabulated function for any value of the argument.[*]

The values of $e^{\omega(r,s;t)}$ at the saddle points were obtained from the tables of $\log G$, using the Bessel's interpolation formula.

In this paper, tables of $\varepsilon\{N^2\}$, $\varepsilon\{N\}$,

$$\sigma^2=\varepsilon\{N^2\}-[\varepsilon\{N\}]^2, \quad \sigma^2/\varepsilon\{N\} \quad \text{and} \quad \sigma^2/([\varepsilon\{N\}]^2-\varepsilon\{N\})$$

are given for various values of y and t. We have also tabulated $G_1(r,s)$ and $G_2(r,s)$ to enable anyone interested to compute $\varepsilon\{N^2(y;t)\}$ for any value of y and t.

---

[*] See for example, Interpolation and Allied Tables (H. M. S. Office).

Table I.

$G_1(r, s)$

r \ s	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8
1.1	−1.9191							
1.2	−1.5890	−1.2365						
1.3	−1.4293	−1.0657	−.89103					
1.4	−1.3221	−.95321	−.77743	−.66437				
1.5	−1.2373	−.86770	−.69267	−.58110	−.49986			
1.6		−.79689	−.62410	−.51486	−.43603	−.37478		
1.7			−.56568	−.45942	−.38338	−.32488	−.27772	
1.8				−.41144	−.33854	−.28292	−.23853	−.20205
1.9					−.29950	−.24692	−.20533	−.17149
2.0						−.21560	−.17683	−.14559
2.1							−.15220	−.12350
2.2								−.10460

$G_1(r, \cdot)$

r \ s	1.9	2.0	2.1	2.2	2.3	2.4	2.5
1.9	−.14345						
2.0	−.11997	−.098757					
2.1	−.10019	−.081107	−.065411				
2.2	−.083507	−.066407	−.052506	−.041219			
2.3	−.069443	−.054189	−.041928	−.032095	−.024254		
2.4		−.044072	−.033299	−.024764	−.018051	−.012824	
2.5			−.026303	−.018920	−.013192	−.0088032	−.0054901
2.6				−.014301	−.0094283	−.0057541	−.0030350
2.7					−.0065474	−.0034792	−.0012552
2.8						−.0018134	.0000011
2.9							.0008576

$G_1(\cdot, s)$

r \ s	2.6	2.7	2.8	2.9	3.0	3.1
2.6	−.0010694					
2.7	.0003100	.0013686				
2.8	.0012412	.0020449	.0025235			
2.9	.0018358	.0024388	.0027661	.0028965		
3.0	.0021830	.0026292	.0028420	.0028911	.0028294	
3.1		.0026773	.0028047	.0027959	.0026972	.0025434
3.2			.0026939	.0026452	.0025237	.0023593
3.3				.0024637	.0023296	.0021625
3.4					.0021293	.0019647
3.5						.0017735

Table II.

$G_2(r, s)$

The starred values of $G_2(r, s)$ in the table are those in the neighbourhood of the points where $\lambda_r + \mu_s - \lambda_{r+s-1}$ or $\lambda_s + \mu_r - \lambda_{r+s-1}$ vanishes. These points do not play any role in the saddle-point integration and so there is no need to include terms involving $e^{-(\lambda_r + \mu_s)t}$ and $e^{-(\lambda_s + \mu_r)t}$ which, as stated earlier, render G_2 finite even at points where $\lambda_r + \mu_s = \lambda_{r+s-1}$ or $\lambda_s + \mu_r = \lambda_{r+s-1}$.

s \ r	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8
1.1	1.1050							
1.2	.12869	.78719						
1.3	−18.563*	.53362	.56617					
1.4	4.6570	.029186	.43086	.39899				
1.5	3.1931	−1.3032	.24178	.31242	.27435			
1.6		−67.833*	−.012904	.21153	.21656	.18464		
1.7			−.40028	.10317	.15639	.14619	.12252	
1.8				−.01266	.098774	.10914	.097407	.080969
1.9					.045543	.075947	.074463	.064860
2.0						.047412	.054726	.050644
2.1							.038367	.038701
2.2								.028954

$G_2(r, s)$

s \ r	1.9	2.0	2.1	2.2	2.3	2.4	2.5
1.9	.053857						
2.0	.043610	.036368					
2.1	.034762	.029833	.025064				
2.2	.027402	.024249	.020838	.017658			
2.3	.021415	.019612	.017240	.014866	.012704		
2.4		.015821	.014239	.012486	.010813	.0093089	
2.5			.011763	.010487	.0091951	.0079960	.0069270
2.6				.0088213	.0078259	.0068676	.0059941
2.7					.0066738	.0059059	.0051892
2.8						.0050893	.0044987
2.9							.0039085
3.0							

$G_2(r, s)$

s \ r	2.6	2.7	2.8	2.9	3.0	3.1
2.6	.0052197					
2.7	.0045438	.0039736				
2.8	.0039584	.0034755	.0030503			
2.9	.0034538	.0030432	.0026786	.0023579		
3.0	.0030197	.0026691	.0023551	.0020771	.0018331	
3.1		.0023456	.0020741	.0018327	.0016196	.0014326
3.2			.0018305	.0016198	.0014331	.0012688
3.3				.0014344	.0012703	.0011255
3.4					.0011285	.0010003
3.5						.00089078

Table III.

$\varepsilon\{N\}$

$[\varepsilon\{N\}]^2$ is given in brackets.

v \ l	2	3	4	5	6	7	8
3	1.98 (3.93)	1.69 (2.87)	1.21 (1.45)	.779 (.606)	.475 (.225)	.280 (.0782)	.160 (.0256)
4	3.90 (15.2)	4.34 (18.8)	3.85 (14.9)	2.98 (8.90)	2.11 (4.47)	1.41 (1.98)	.897 (.805)
5	6.87 (47.3)	9.47 (89.8)	10.2 (103)	9.27 (86.0)	7.55 (56.9)	5.67 (32.1)	4.01 (16.1)
6	11.2 (125)	18.6 (345)	23.4 (549)	24.7 (611)	22.9 (524)	19.3 (372)	15.1 (228)
7		33.3 (1110)	48.8 (2380)	58.8 (3450)	61.3 (3760)	57.5 (3310)	49.6 (2460)
8		56.4 (3180)	94.2 (8870)	128 (1.63×10^4)	149 (2.22×10^4)	154 (2.38×10^4)	146 (2.13×10^4)
9			171 (2.92×10^4)	259 (6.71×10^4)	334 (1.12×10^5)	381 (1.45×10^5)	393 (1.54×10^5)
10			296 (8.73×10^4)	496 (2.46×10^5)	703 (4.94×10^5)	874 (7.64×10^5)	979 (9.58×10^5)
12				1580 (2.51×10^6)	2660 (7.10×10^6)	3890 (1.51×10^7)	5070 (2.57×10^7)
14					8620 (7.42×10^7)	14500 (2.11×10^8)	21700 (4.70×10^8)
16						47400 (2.25×10^9)	80100 (6.41×10^9)

Table IV.

$\varepsilon\{N^2\}$

σ^2 is given in brackets.

v \ l	2	3	4	5	6	7	8
3	6.48 (2.55)	4.49 (1.62)	2.57 (1.12)	1.37 (.76)	.740 (.515)	.400 (.322)	.216 (.190)
4	21.6 (6.4)	25.4 (6.6)	18.6 (3.7)	12.1 (3.2)	6.89 (2.42)	3.76 (1.78)	2.05 (1.24)
5	63.1 (15.8)	114 (24)	126 (23)	101 (15)	67.9 (11.0)	40.6 (8.5)	22.8 (6.7)
6	163 (38)	424 (79)	650 (101)	698 (87)	584 (60)	414 (42)	261 (33)
7		1360 (250)	2800 (420)	3910 (460)	4140 (380)	3570 (260)	2640 (180)
8		3900 (720)	1.04×10^4 $(.15 \times 10^4)$	1.86×10^4 $(.23 \times 10^4)$	2.45×10^4 $(.23 \times 10^4)$	2.57×10^4 $(.19 \times 10^4)$	2.26×10^4 $(.13 \times 10^4)$
9			3.45×10^4 $(.53 \times 10^4)$	7.67×10^4 $(.96 \times 10^4)$	1.24×10^5 $(.12 \times 10^5)$	1.57×10^5 $(.12 \times 10^5)$	1.64×10^5 $(.10 \times 10^5)$
10			1.03×10^5 $(.16 \times 10^5)$	2.82×10^5 $(.36 \times 10^5)$	5.53×10^5 $(.59 \times 10^5)$	8.36×10^5 $(.72 \times 10^5)$	1.03×10^6 $(.07 \times 10^6)$
12				2.91×10^6 $(.40 \times 10^6)$	8.05×10^6 $(.95 \times 10^6)$	1.68×10^7 $(.17 \times 10^7)$	$2,80 \times 10^7$ $(.23 \times 10^7)$
14					8.53×10^7 (1.11×10^7)	2.37×10^8 $(.26 \times 10^8)$	5.19×10^8 $(.49 \times 10^8)$
16						2.56×10^9 $(.31 \times 10^9)$	7.16×10^9 $(.75 \times 10^9)$

Table V.

$$\sigma^2 \; ([\varepsilon\{.V\}]^2 - \varepsilon\{.V\})$$

$\sigma^2/\varepsilon\{.V\}$ is given in brackets

y \ t	2	3	4	5	6	7	8
3	1.3 (1.3)	1.4 (.96)	4.7 (.93)	−4.4 (.98)	−2.1 (1.1)	−1.6 (1.2)	−1.4 (1.2)
4	.57 (1.6)	.46 (1.5)	.34 (.96)	.54 (1.1)	1.0 (1.1)	3.1 (1.3)	−1.3 (1.4)
5	.39 (2.3)	.30 (2.5)	.25 (2.2)	.20 (1.6)	.22 (1.5)	.32 (1.5)	.55 (1.7)
6	.33 (3.4)	.24 (4.2)	.19 (4.3)	.15 (3.5)	.12 (2.6)	.12 (2.2)	.15 (2.2)
7		.23 (7.5)	.18 (8.6)	.14 (7.8)	.10 (6.2)	.080 (4.5)	.075 (3.6)
8		.23 (13)	.17 (16)	.14 (18)	.10 (15)	.081 (12)	.061 (8.9)
9			.18 (31)	.14 (37)	.11 (36)	.083 (31)	.064 (25)
10			.18 (54)	.15 (73)	.12 (84)	.094 (82)	.073 (72)
12				.16 (250)	.13 (360)	.11 (440)	.089 (450)
14					.15 (1300)	.12 (1800)	.10 (2300)
16						.14 (6500)	.12 (9400)

We have calculated $\varepsilon\{V(y; t)\}$ using the tables of Bhabha and Chakrabarthy**. It is to be noted that our values for $\varepsilon\{V(y; t)\}$ differ from those given in Heitler's " Quantum theory of radiation " (H.2).

Discussion of results

In discussing our results from the standpoint of the general theory of multiplicative stochastic processes we meet with one serious difficulty : the total cross-section for radiation by an electron is infinite if we use the exact Bethe-Heitler expression for $R^{(1)}$ given by (22). Consequently, all the moments of the *total* number of particles (i. e. $N(y; t)$ as y tends to infinity or in other words, the number of particles above zero energy) become infinite. This is contrary to physical facts. So in (D. 6) Ramakrishnan *assumed* the total cross-section to be finite and obtained an approximately Furry distribution for the total number of electrons.

The above difficulty does not arise if we consider the number of particles above a certain energy i. e. for finite y (which can be chosen as large as we please). Fortunately, it is the case of finite y that is of physical significance and the calculation of the moments of the total number is only of academic interest. However, the question arises :

** We have in fact used the tables of Leonie Janossy and Messel (Proc. Roy. Irish Acad. Sci. **54A** (1951), 217) who have extended those of Bhabha and Chakrabarthy to more decimal places,

is there any essential difference from a stochastic point of view, between the case where the total cross-section is finite and where it is infinite? In a later paper, one of us (R) proposes to discuss two simplified models of multiplicative processes, one in which the total cross-secton is finite and the other in which it is infinite. It is hoped that a comparison between these two models may reveal any marked differences in their stochastic features which are obscured in more complicated processes by the very complexity of the problem.

In the present paper we have tabulated $\sigma^2/\varepsilon\{N\}$ and $\sigma^2/([\varepsilon\{N\}]^2-\varepsilon\{N\})$ for convenience of comparison with the Poisson and Furry distributions. It should be noted that $\sigma^2/\varepsilon\{N\}$ and $\sigma^2/([\varepsilon\{N\}]^2-\varepsilon\{N\})$ are complicated functions of y and t as is obvious from the fact that we are not able to invert the double complex integral (43). An examination of the tables reveals the following general features of the process.

1. For a given y the maxima for $\varepsilon\{N\}, \varepsilon\{N^2\}, \sigma^2, \sigma^2/\varepsilon\{N\}$ and $\sigma^2/([\varepsilon\{N\}]^2-\varepsilon\{N\})$ are all different as expected.

2. Comparison with the Furry distribution is meaningful only if the number of particles is more than, say, 10. The negative values for $\sigma^2/([\varepsilon\{N\}]^2-\varepsilon\{N\})$ correspond to the negative values for $([\varepsilon\{N\}]^2-\varepsilon\{N\})$ i. e. when $\varepsilon\{N\}<1$. For very large values of y but before the shower maximum the deviation is comparable to (about .1 to .2 times) the Furry deviation.

3. Comparing with the Poisson distribution, we find that $\sigma^2/\varepsilon\{N\}$ increases with y for a given value of t. At the shower maximum the deviation is many times the Poisson value – a result at variance with that of Scott and Uhlenbeck and Janossy and Messel. For example, for $y=8$ the maximum occurs between $t=6$ and $t=8$ and the deviation is about 12 times the Poisson value or .1 times the Furry deviation.

While we do not hazard a detailed criticism of the results of Scott and Uhlenbeck or Janossy and Messel we here attempt to indicate the reasons for such a discrepancy.

(i) It was pointed out by Bhabha and Ramakrishnan that Scott and Uhlenbeck, in the course of their calculations erroneously omitted certain terms as negligible.

(ii) The calculations of Janossy and Messel were based on the equation for the second moment obtained from the fundamental integral equations of Janossy. This equation did not yield an explicit Mellin's transform solution for $\varepsilon\{N^2\}$ (see equations (22), (23) and (24) of their paper) and this is as should be expected since they were dealing with the direct Mellin's transform of $\varepsilon\{N^2\}$ with respect to y while Bhabha and Ramakrishnan obtained $\varepsilon\{N^2\}$ as the inverse of a two-fold Mellin's transform.

(iii) *There seems to be no reason, from a logical and stochastic point of view, to expect a Poisson distribution at the shower maximum.* This expectation by previous authors was probably based on the assumption that in a 'stationary state,' particles behave as if they are independent and hence the number of particles $N(y; t)$ has a Poisson distribution at thicknesses where $\varepsilon\{N(y; t)\}$ is a maximum. But $\varepsilon\{N(y; t)\}_{max}$ does in no way define a 'stationary state' in a stochastic sense. In the theory of stochastic processes we call a 'state' stationary only if the differential coefficients (with respect to t) of *all* orders of *all* the moments vanish.

In conclusion we wish to assert that our calculations are as accurate as can be obtained by the use of the double saddle-point formula. Our equations for the product density of degree two are a natural sequel to the equations of Bhabha and Chakrabarthy for the mean i. e. product density of degree one, since both are 'last collision' diffusion equations – as contrasted with the 'first collision' equations of Janossy and Messel.

Appendix

Here we point out the mistakes occurring in the paper of Bhabha and Ramakrishnan*.

1.   On page 144 of that paper it was stated (see equation (5)) that in the absence of the last terms in equations (4a) and (4b), equation (4) would clearly be satisfied by

$$f_2(E_1, E_2;\ t) = f_1(E_1;\ t)f_1(E_2;\ t),$$

$$fg_{1,1}(E_1, E_2;\ t) = f_1(E_1;\ t)g_1(E_2;\ t),$$

$$fg_{1,1}(E_2, E_1;\ t) = f_1(E_2;\ t)g_1(E_1;\ t),$$

$$g_2(E_1, E_2;\ t) = g_1(E_1;\ t)g_1(E_2;\ t),$$

The correct statement should be

$$f_2(E_1, E_2;\ t) = k_1 f_1(E_1;\ t)f_1(E_2;\ t),$$

$$fg_{1,1}(E_1, E_2;\ t) = k_2 f_1(E_1;\ t)g_1(E_2;\ t),$$

$$fg_{1,1}(E_2, E_1;\ t) = k_3 f_1(E_2;\ t)g_1(E_1;\ t),$$

$$g_2(E_1, E_2;\ t) = k_4 g_1(E_1;\ t)g_1(E_2;\ t),$$

where k_1, k_2, k_3, k_4 are constants. The initial condition in the case of a shower initiated by a single electron of energy E_0 is

$$f_1(E;\ 0) = \delta(E - E_0)\quad g_1(E;\ 0) = 0 \qquad \text{at } t = 0,$$

$$f_2(E_1, E_2;\ 0) = f g_{1,1}(E_1, E_2;\ 0) = g_2(E_1, E_2;\ 0) = 0 \quad \text{for all } E_1, E_2.$$

This follows directly from the definition of product densities. Thus $k_1 = k_2 = k_3 = k_4 = 0$. *Fortunately this error was not carried over into the calculations* since $\phi^{(2)}(r, s;\ t)$ the vector representing the transforms of the product densities was expressed as an integral 0 to t according to (25) of that paper. This means that at $t = 0$, the vector is null.

2.   The expression for $R(E', E)$ as given by (6a) is not correct. It should read

$$R(E', E) = \left\{ \frac{E - E'}{E} - \left(\frac{4}{3} + a \right)\left(1 - \frac{E}{E - E'} \right) \right\} \frac{1}{E} = R^{(1)}(E', E)$$

since in our definition of $R^{(1)}(E', E)$, E' represent the energy of the electron after radiation and not that of the photon. But this error was not carried over into the calcu-

---

lations since the correct expression for A_r, A_s was written down.

3. The vector v' of equation (35) of that paper has to be multiplied by a factor $\dfrac{1}{(\mu_s - \lambda_s)(\mu_r - \lambda_r)}$. This factor was omitted by mistake.

4. On page 150, the same expression is numbered as (36) and (37).

5. In eq. (40) there is no need to separate out f_0 and f''. $f_0 + f''$ can be written as f' and so in (41), we write f' and v' instead of f'' and v''. Using the usual approximation for all but small t we get (47) of the present paper.

References

A. Cascade theory (Mean number)

1. H. J. Bhabha and W. Heitler, Proc. Roy. Soc. A **159** (1937), 432.
2. J. F. Carlson and J. R. Oppenheimer, Phys. Rev. **51** (1937), 220.
3. L. Landau and G. Rumer, Proc. Roy. Soc. A **166** (1938), 267.
4. H. S. Snyder, Phys. Rev. **53** (1938), 960.
5. R. Serber, Phys Rev. **54** (1938), 317.
6. D. Iwanenko and A. Sokolow, Phys. Rev. **53** (1938), 910.
7. W. Heisenberg and H. Euler, Ergeb. exact, Naturw. **17** (1938), 1.
8. I. Tamm and S. Belenky, Jou. Phys. U. S. S. R. **1** (1932), 177.
9. L. W. Nordheim and M. H. Hebb, Phys. Rev. **56** (1939), 494
10. B. Rossi and K. Greisen, Rev. Mod. Phys. **13** (1941), 240.
11. H. J. Bhabha and S. K. Chakrabarthy, Proc. Roy. Soc. A **181** (1943), 267.
12. L. Janossy, Comm. Dublin Inst. Adv. Studies A **4** (1947).
13. I. Tamm and S. Belenky, Phys. Rev. **70** (1946), 660.
14. H. J. Bhabha and S. K. Chakrabarthy, Phys. Rev. **74** (1948), 1352.
15. J. Nishimura and K. Kamata, Prog. Theor. Phys. (Japan) **7** (1952), 185.
16. H. S. Snyder Phys. Rev. **80** (1950), 611.

B. Cascade theory (Fluctuation problem with simplified models)

1. W. H. Furry, Phys. Rev. **52** (1937), 569.
2. N. Arley, Proc. Roy. Soc. (London) A **168** (1938), 519.
3. H. Euler, Z. Physik **110** (1938), 450.
4. N. Arley, *On the theory of stochastic processes* John Wiley, New York, (1949).
5. A. Nordseik, W. E. Lamb and G. E. Uhlenbeck, Physica **7** (1940), 344.
6. W. T. Scott and G. E. Uhlenbeck, Phys. Rev. **62** (1942), 497.

C. Method of product densities

1. Alladi Ramakrishnan, Proc. Camb. Phil. Soc. **46** (1950), 595.
2. H. J. Bhabha, Proc. Roy. Soc. (London) A **202** (1950), 301.
3. D. G. Kendall, J. R. Statist. Soc. (B) **11** (1949), 230.
4. H. J. Bhabha and A. Ramakrishnan, Proc. Ind. Acad. Sci. **32** (1950), 141.

D. Method of product densities

1. R. Bellmann and T. E. Harris, Proc. Nat. Acad. Sci. (Washington) **34** (1948), 601.
2. C. Palm, Ericsson Technics No. 44 (1943), 1–189.
3. L. Janossy, Proc. Phys. Soc. (London) A **64** (1951), 807.

4. L. Janossy and H. Messel, Proc. Phys. A **93** (1950), 1101.
5. M. S. Bartlett and D. G. Kendall, Proc. Camb. Phil. Soc. **47** (1950), 65.
6. Alladi Ramakrishnan, J. R. Statist. Soc. (B) **13** (1951), 131.
7. M. S. Bartlett, Proc. Camb. Phil. Soc. **47** (1951), 821.

E. Method of Janossy functions

1. L. Janossy, Proc. Roy. Irish Acad. A **53** (1950), 181.

F. Further contributions using the above methods

1. H. Messel and J. W. Gardner, Phys. Rev. **84** (1951), 1256.
2. H. Messel, Phys. Rev. **83** (1951), 21 ; **83** (1951), 26.
3. H. Messel, Proc. Phys. Soc. (London) A **64** (1951), 807.
4. H. Messel, Proc. Roy. Irish Acad. A **54** (1951a), 125.
5. H. Messel, Commun. Dublin Inst. for Adv. Studies Series A **7** (1951b).
6. H. Messel, Proc. Phys. Soc. (London) A **65** (1952), 465.
7. H. Messel and R. B. Potts, Proc. Phys. Soc. (London) A **65** (1952), 473.
8. H. Messel and R. B. Potts, Phys. Rev. **86** (1952), 847.
9. H. S. Green and H. Messel, Proc. Camb. Phil. Soc. to be published.
10. H. Messel and R. B. Potts, Proc. Phys. Soc. A **65** (1952), 473.
11. H. Messel and R. B. Potts, Phys Rev. **87** (1952), 738.
12. H. Messel and J. W. Gardner, Phys. Rev. **86** (1952), 808.
13. H. Messel and H. S. Green, Phys. Rev. **87** (1952), 738.

G. Articles comparing the methods or in the nature of reviews

1. Alladi Ramakrishnan, Proc. Camb. Phil. Soc. **48** (1952), 451.
2. H. Messel and R. B. Potts, Proc. Phys. Soc. (London) A **65** (1952), 854.
3. W. T. Scott, Phy. Rev. **82** (1951), 893.
4. L. Janossy, Acta. Physica (Budapest) **11** (1952), 304.

H. Books

1. I. Janossy, *Cosmic rays*, 2nd Ed. (Oxford, 1944).
2. W. Heitler, *Quantum theory of radiation*, 2nd Ed. Oxford, (1944).
3. H. Messel, *Progress in cosmic ray physics* North Holland Publishing Co., Amsterdam, (1952) Vol. II.
4. W. Feller, *Introduction to probability theory* John Wiley, (1950).

Letters to the Editor

Nucleon Isobars and Nuclear Forces in Low Energy Region

Takashi Kikuta

Department of Physics, University of Tokyo

December 4, 1953

The existence of the nucleon isobars has been considered to be probable from the experimental data[1] and also from the theoretical predictions.[2] The purpose of this note is to estimate roughly various effects of nucleon isobars for nuclear forces in low energy region.

To do this, we borrow " Pauli-Kusaka " potential[3] which is derived from pv coupling interaction of ps symmetrical meson theory in the strong coupling limit, and solve the problem by imposing the condition that nucleon isobar states are present only when two nucleons approach each other within the range of nuclear forces. For simplicity, the radial parts of interaction potentials are assumed to be square well (both central and tensor force range are equal), and their range and depth are adjusted to agree with constants* of the singlet low energy scattering (scatt. length, effective range) and the deuteron state (binding energy, electric quadrupole moment) by omitting isobar waves, as follows[4]

range of potential; 2.65×10^{-13}cm,

central force ; $V = 2.3$: $r < 1$, $= 0$: $r > 1$,

tensor force ; W 2.3 : r 1, 0 : r 1.

Since the various isobar states are present, the fundamental equations for usual spin singlet and triplet states are :

for singlet :

$$(d^2/dr^2 + k^2 + V^-)u_0^{(1)} - \sqrt{5}\, V^- u_2^{(1)}$$

$$- 4W v_1^{(5)} + \sqrt{5}\, W v_2^{(5)} = 0,$$

* Pauli-Kusaka's potential gives same central potential for both spin singlet and triplet states. Therefore there is no ambiguity for agreement with the effective range of triplet state.

$$(d^2/dr^2 + k^2 - \varepsilon - 6/r^2 + V^-/2 - 5W^-/2)v_1^{(5)}$$

$$- 4V^- u_0^{(1)} + \sqrt{4/5}\, W u_2^{(5)} \doteq 0. \quad \text{etc}$$

(omitting higher waves).

for triplet :

$$(d^2/dr^2 + k^2 + V^-)u_0^{(3)} + 2\sqrt{2}\, W v_0^{(3)} - \sqrt{5}\, V u_2^{(3)}$$

$$- \sqrt{2/5}\, W v_2^{(3)} + 3\sqrt{7/5}\, W v_2^{(7)} = 0, \quad \text{etc.}$$

where u and v are S and D-wave functions respectively, and the superfixes in the brackets show the spin multiplicity. Indices 0, 1, 2 express the ground state (usual two nucleon state), first excited state with excitation energy ε, second excited state (2ε) respectively. In practical calculations, the first excitation energy ε is taken as 300 Mev.[5] Rewriting the above equations in the following form

$$(d^2/dr^2 + k^2 + V^- + V_2^{-1} + V_1^{-5} + V_2^{-5})u_0^{(1)} = 0,$$

$$V_2^{-1} = -\sqrt{5}\, V u_2^{(1)}/u_0^{(1)}, \quad \text{etc,}$$

$$(d^2/dr^2 + k^2 + V^- + V_0^{-3} + V_2^{-3} + V_2^{-7} + V_2^{-7})u_0^{(3)} = 0,$$

r si	0.0	0.5	1.0	1.4
$u_0^{(1)}$	$1.55r$	0.687	0.999	1.033
$u_2^{(1)}$	$-0.07r$	-0.032	-0.024	-0.000
$v_1^{(5)}$	$O(r^3)$	-0.076	-0.072	-0.004
v_2	$O(r^3)$	0.027	-0.022	0.000
V_2^{-1}/V^-	0.106	0.105	0.053	0
V_1^{-5}/V	0.000	0.443	0.287	0
V_2^{-5}/V	0.000	0.086	0.050	0

r tr	0.0	0.5	0.8	1.0
$u_0^{(3)}$	$1.87r$	0.804	0.997	0.956
$u_2^{(3)}$	$-0.09r$	-0.039	-0.046	-0.025
$v_2^{(3)}$	$O(r^3)$	-0.013	-0.017	-0.010
$v_2^{(7)}$	$O(r^3)$	0.047	0.062	0.034
V_2^{3}/V^-	0.113	0.113	0.107	0.060
V_2^{-7}/V	0.000	0.015	0.016	0.009
V_2^{-7}/V	0.000	0.202	0.215	0.124

Table 1. Probability amplitudes of isobar states (with the ground S-states) and the ratios of effective potentials and the central potential for the ground state.

where V_2', etc., may be considered as the effective potentials which are added to the ground state potential on account of the presence of various isobar states, they are solved approximately at zero energy, assuming that V/ε and W/ε are small quantities, and the results are summarized in Table 1.

(Conclusions)

(1). The probability amplitude for the fixed isobar wave is inversely proportional to its excitation energy approximately.

(2). The amplitude for the isobar wave is at most 10% to the S-wave of the ground state within the nuclear forces, however, the depth of the largest effective potential (V_1^5) reaches 1/3 of the depth of the ground state central potential.

(3). The effective potentials resulting from the isobar states always act as attractive forces.

(4). The effects for the spin singlet is larger than that of the spin triplet state. However we can rearrange the depths and the ranges of potentials including the isobar effects so as to agree with the values of parameters for low energy nuclear forces.

In this case the depth of central potential becomes to be smaller considerablly, while the tensor's is not so changed.

(5). Since the probability amplitudes of nucleon isobars in outside region of nuclear forces damp strongly, the effective potentials vanish at the region of tails of nuclear forces. Moreover since the D-states predominate among other isobars, the contribution of effective potentials is remarkable in the vicinity of the force range.

(6). The effects for the effective range, electric quadrupole moment and D-state probability of deuteron seem to be small, for these quantities contain the effects of isobars in form of square of probability amplitudes as follows:

$$r_{0s} = 2\int_0^\infty (\psi_s'^2 - u_0^{(1)2} - u_2^{(1)2} - v_1^{(5)2} - v_2^{(5)2})\,dr,{}^*$$

$$r_{0t} = 2\int_0^\infty (\psi_t'^2 - u_0^{(3)2} - v_0^{(3)2} - u_2^{(3)2} - v_2^{(3)2} - v_2^{(7)2})\,dr.$$

where

$$\psi_s = \lim_{k\to 0} (\cos kr + \cot \delta_s \sin kr) \simeq 1 - a_s r,$$

* The correction resulting from the Coulomb force in proton-proton system is small. (c.f. reference 6).

and (r_0 is the range of nuclear forces),

$$u_0^{(1)} = \phi_s,\quad u_2 = v_1 = v_2 = 0,$$

$$u_0^{(3)} = \lim_{k\to 0}\cos\varepsilon\,\psi_t = 1 - a_t r,$$
$$v_0^{(3)} = \lim_{k\to 0}(-\sin\varepsilon)\psi_t = 3q/r^2,\quad ;\ \text{for}\ r\gg r_0.$$

$$\tan\varepsilon = qk^2 + \cdots\cdots$$

For the deuteron state, putting

$$\int_0^\infty (u_0^{(3)2} + v_0^{(3)2} + u_2^{(3)2} + v_2^{(3)2} + v_2^{(7)2})\,dr = 1,$$

it follows that

$$Q = \int_0^\infty [\sqrt{2}/10\cdot(u_0^{(3)}v_0^{(3)} + u_2^{(3)}v_2^{(3)})$$
$$- 1/20\cdot(v_0^{(3)2} + v_2^{(3)2}) - 1/70\cdot v_2^{(7)2}]r^2 dr,$$

$$P_{\text{D-state}} = \int_0^\infty (v_0^{(3)2} + v_2^{(3)2} + v_2^{(7)2})\,dr.$$

Probably the indirect effects through the ground state by the presence of isobars may be rather larger than the direct one.**

(7). It is possible, however, that the effects of isobars are overestimated in the present treatment by taking the matrix elements from the strong coupling limit, even though the excitation energy of the lowest isobar is put so high as 300 Mev.

The author wishes to express his cordial thanks to Professors M. Taketani, T. Yamanouchi, M. Sasaki, H. Fukuda, Dr. H. Miyazawa, Messrs. A. Fujii, J. Iwadare and other researchers of nuclear forces for their interests and valuable discussions.

1) Fujimoto and Miyazawa, Prog. Theor. Phys. 5 (1950), 1052.
Brueckner and Case, Phys. Rev. 83 (1951), 1141.
K. A. Brueckner, Phys. Rev. 86 (1952), 106.

** G. Breit and M. C. Yovits (reference 7) had discussed the range of nuclear forces rather phenomenologically assuming isobar of S-state. Since the effective potentials from the isobar waves damp strongly in the tail of nuclear force, the effective potentials from S-state isobars shorten the range of nuclear forces. However, according to the strong coupling theory the D-state isobars are so predominant that the tendency of shortening the range is canceled out by the centrifugal force. For the deuteron state, since the wave functions of the ground state spread far away out of the force range, the effects of isobars for Q and P may be still less.

Anderson, Fermi, Martin and Nagle. Phys. Rev. **91** (1953), 155.

Cocconi and Silverman, Phys. Rev. **83** (1952), 1230.

Anderson et al., Phys. Rev. **86** (1952), 413.

Matsuyama and Miyazawa, Prog. Theor. Phys. **9** (1953), 492.

2) G. Wentzel, Rev. Mod. Phys. **19** (1947), 1.
    S. Tomonaga, Prog. Theor. Phys. **2** (1947), 2.
    Maki, Sato and Tomonaga, Prog. Theor. Phys. **9** (1953), 607.

3) Pauli and Kusaka, Phys. Rev. **63** (1943), 400.
    M. Fierz, Helv. Phys. Acta **17** (1944), 181 and **18** (1945), 158.
    F. Villars, Helv. Phys. Acta, **19** (1946), 323.

4) Christian and Hart, Phys. Rev **77** (1950), 441, Fig. 11, Fig. 13.

5) See reference 1).

6) H. A. Bethe, Phys. Rev. **76** (1949), 38.

7) Breit and Yovits, Phys. Rev. **81** (1951), 416.

Fermi's Theory of the Generation of Pions and the Partition Theory of Numbers

F. C. Auluck and D. S. Kothari

University of Delhi, India

December 9, 1953

In recent years the partition function $p_k(n)$ has received considerable attention: $p_k(n)$ represents the number of distinct ways of partitioning an integer n into exactly k integral parts $(a_1, a_2, \ldots\ldots a_k)$, where $n = a_1 + a_2 \ldots\ldots + a_k$. If k_0 denotes the value k for which $p_k(n)$ is maximum, then Szekeres[1] has obtained for k_0 the expression

$$k_0 = \left(\frac{6n}{\pi^2}\right)^{1/2} L + \frac{9}{\pi^2}\left(1 + L - \frac{1}{6}L^2\right) - \frac{1}{2}$$
$$+ O(n^{-1/2}\log^4 n), \qquad (1)$$

where $\qquad L = \log(6n/\pi^2)^{1/2}$,

This confirms the conjecture of Auluck, Chowla and Gupta[2] regarding the uniqueness of the maximum for $p_k(n)$. If $\bar{k}$ denotes the mean value of k defined according to

$$\bar{k} = \sum_{k=1}^{n} k\, p_k(n) \Big/ \sum_{k=1}^{n} p_k(n), \qquad (2)$$

then, it can be shown that (Auluck: unpublished result) $\bar{k}$ is

$$\bar{k} = \left(\frac{6n}{\pi^2}\right)^{1/2}(L + C) + \frac{3}{2\pi^2}C + \frac{1}{4}$$
$$+ O(n^{-1/2}\log n), \qquad (3)$$

where C is Euler's constant. (The difference between k_0 and $\bar{k}$ may be noted).

The present note is concerned with the application of the foregoing results to the problem of the production of pions in high-energy nucleon collisions. Consider a nucleon of high-energy W' colliding with another nucleon at rest. In the 'laboratory system' the total energy and momentum of the two nucleons is respectively $M(\gamma'+1)$ and $M(\gamma'^2-1)^{1/2}$ where $\gamma' = W'/M$; M being the nucleon rest-mass: We take the velocity of light as unity. Thus, in the coordinate system in which the total momentum is zero, the energy W is given by

$$W \equiv 2\gamma M = [M^2(\gamma'+1)^2 - M^2(\gamma'^2-1)]^{1/2}$$
$$= M(2\gamma'+2)^{1/2}. \qquad (4)$$

It is this energy W which is available for the production of pions. If μ denotes the rest-mass of a pion and ν stands for the number W/μ, then we assume that the statistical problem of the production of pions can be conceived as essentially equivalent to a partitioning of ν into integral parts. On this assumption, the most probable number of pions produced will be the value of k for which $p_k(\nu)$ is maximum. Hence, if N denotes the most probable number of pions produced in the collision process, we have

$$N = k_0 \sim (\sqrt{6}/\pi^2 \cdot \log\sqrt{6\nu/\pi^2})\nu^{1/2}. \qquad (5)$$

If $\bar{N}$ denotes the average number produced, we have from equation (3)

$$\bar{N} = \bar{k} \sim (6/\pi^2)^{1/2}(\log\sqrt{6\nu/\pi^2} + C)\nu^{1/2}. \qquad (6)$$

The above relations are to be compared with the expression given by Fermi[3] for the most probable number of pions:

$$N \frac{(8\pi)^{1/2}\zeta(3)}{(6\pi\zeta(4))^{3/4}}\left(\frac{M}{\mu}\right)^{3/4}\left(\frac{W'}{M}\right)^{1/2} \sim 1.0\nu^{1/2}. \qquad (7)$$

(It will be noted that Fermi's expression as well as the relations obtained here do not involve $\hbar$). It will be seen that for reasonable values of ν, the agreement between the Fermi-expression (7) and

the equation (5) is satisfactory. In view of the considerable interest in the Fermi Theory, it is not altogether without interest to see that essentially the same result follows on the above picture of the excitation of " pion oscillators ".

1) G. Szekeres, Quart. Jour. of Math. Oxford Series **4** (1953), 96.
2) F. C. Auluk, S. Chowla and M. Gupta, Jour. Ind. Math. Soc. **6** (1942), 105.
3) E. Fermi, Prog. Theor. Phys. **5** (1950), 570.

The Formal Theory of Error Estimation of Rearrangement Collisions

Tatuya Sasakawa

Department of Physics, Kyoto University

December 17, 1953

Recently, the theory of rearrangement collisions was treated in a complete general way along the line of Lippmann and Schwinger's theory[1] by Fujimoto, Hayakawa and Nishijima.[2] On the other hand, a generalization of the impulse approximation with special attention given to multiple scattering effects was formulated in a rigorous way by Chew and Goldberger[3].

In the present paper, by a similar procedure as Chew and Goldberger it is shown how the error of rearrangement collision can be expressed as the sum of errors associated with the entrance channel plus ones associated with the outgoing channel plus the wave function used.

The total Hamiltonian is decomposed into two ways

$$H = H_a + H_a' = H_b + H_b', \qquad (1)$$

where a and b refer the entrance channel and the outgoing channel respectively.

T-matrix is given by[2]

$$T_{ba} = (\Psi_b^{(-)}, H_a' \, \Phi_a) = (\Phi_b, H_b' \, \Psi_a^{(+)})$$

$$= (\Phi_b, H_a' \, \Phi_a) + \left(\Phi_b, H_b' \frac{1}{E_a + i\eta - H} H_a' \Phi_a \right)$$

$$= (\Phi_b, H_b' \, \Phi_a) + \left(\Phi_b, H_b' \frac{1}{E_a + i\eta - H} H_a' \Phi_a \right), \qquad (2)$$

where $(H_a - E_a)\Phi_a = 0$ and $(H_b - E_b)\Phi_b = 0$. (3)

Defining the notation Λ_{ba} and Ω_a as

$$\Lambda_{ba} = H_b' \Omega_a = H_b' + H_b' \frac{1}{E_a + i\eta - H} H_a', \qquad (4)$$

(2) can also be expressed simply as

$$T_{ba} = (\Phi_b, \Lambda_{ba} \, \Phi_a). \qquad (5)$$

Suppose we make use of an approximate Hamiltonian instead of true Hamiltonian (1), for the sake of simplifying calculation or from any other reason,

$$\mathfrak{H} = H_i + H_i'. \qquad (6)$$

In this case T-matrix is given by

$$\mathfrak{T}_{ba} = (\Phi_{bi}, \Lambda_{bi, ai} \Phi_{ai}), \qquad (7)$$

where

$$\Lambda_{bi, ai} = H'_{bi} + H'_{bi} \frac{1}{E_{ai} + i\eta - \mathfrak{H}} H'_{ai} = H'_{bi} \Omega_{ai}$$

as before. (8)

Errors originated in the choice of this approximate Hamiltonian (6) is

Errors
$$= T_{ba} - \mathfrak{T}_{ba} = (\Phi_b, \Lambda_{ba}\Phi_a) - (\Phi_{bi}, \Lambda_{bi, ai}\Phi_{ai})$$

$$= \{(\Phi_b, \Lambda_{ba}\Phi_a) - (\Phi_b, \Lambda_{bi, ai}\Phi_a)\}$$

$$+ \{(\Phi_b, \Lambda_{bi, ai}\Phi_a) - (\Phi_{bi}, \Lambda_{bi, ai}\Phi_{ai})\}. \qquad (9)$$

The last pair of terms is the error arisen from the approximate wave function. The difference of the first pair of terms of Eq. (9) is calculated as

$$\Lambda_{ba} - \Lambda_{bi, ai} = \left\{ \left(H_b' + H_b' \frac{1}{E_a + i\eta - H} H_a' \right) \right.$$

$$\left. - \left(H'_{bi} + H'_{bi} \frac{1}{E_b + i\eta - H} H'_{ai} \right) \right\}$$

$$+ \left\{ \left(H'_{bi} + H'_{bi} \frac{1}{E_a + i\eta - H} H'_{ai} \right) \right.$$

$$\left. - \left(H'_{bi} + H'_{bi} \frac{1}{E_{ai} + i\eta - \mathfrak{H}} H'_{ai} \right) \right\}$$

$$= (H'_b - H'_{bi}) + (H'_b - H'_{bi}) \frac{1}{E_a + i\eta - H} H_a'$$

$$+ H'_{bi} \frac{1}{E_a + i\eta - H} (H_a' - H'_{ai})$$

$$+ H'_{bi} \frac{1}{E_a + i\eta - H} (E_{ai} - E_a + H - \mathfrak{H})$$

$$\times \frac{1}{E_{ai} + i\eta - \mathfrak{H}} H'_{ai}$$

$$= (H_b' - H_b i')\Omega_a + H_b i' \frac{1}{E_a + i\eta - H}$$

$$\times \{[H_a{}^0 - H_{ai}{}^0, \Omega_{ai}] + (H_a' - H_{ai}')\Omega_{ai}\}. \quad (10)$$

Thus

$$\text{Errors} = (\Phi_b, \Delta_b' \Psi_a{}^{(+)}) + \left(\Phi_b, H'_b i \frac{1}{E_a + i\eta - H}\right.$$

$$\{[\Delta_a{}^0, \Omega_{ai}] + \Delta'_a \Omega_{ai}\}\Phi_a \Big)$$

$$+ (\{\Phi_b, H'_b i \Omega_{ai}\Phi_a\} - (\Phi_b i, H'_b i \Omega_{ai}\Phi_{ai})\},$$

where

$$\Delta'_b = H'_b - H'_b i,$$

$$\Delta_a{}^0 = H_a{}^0 - H^0{}_{ai}$$

and

$$\Delta'_a = H_a' - H'_{ai}. \quad (11)$$

The significance of these terms is suggested by the observation that when Δ'_b vanishes while other terms remain, and so on.

Finally in general,

Errors = (Errors associated with the entrance channel)

+ (Errors associated with the outgoing channel)

+ (Errors associated with the wave function used).

$$(12)$$

A similar procedure can be developed for exchange reaction, in which we are concerned with K-matrix[1] instead of T-matrix.

As a simple example, though not the best one, of the above formulation, let us take the theory of stripping reaction, which errors is already argued by Gerjuoy[4] and Fujimoto[5].

The Hamiltonian of the total system is

$$H = T_P + T_N + H_\xi + V_{NP} + V_{N\xi} + V_{P\xi}, \quad (13)$$

where ξ, N and P indicate the target nucleus, neutron and proton respectively. Gerjuoy expressed the solution satisfies

$$(H - E)\Psi = 0$$

as

$$\Psi = \psi_D + \frac{1}{E - H_0}(V_N - V_{NP})\psi_D$$

$$+ \frac{1}{E - H_0}(V_P + V_{NP})\Psi, \quad (14)$$

where ψ_D is expressed in terms of the solution satisfies

$$(H_0 - E)\psi_0 = 0, \quad H_0 = T_N + T_P + V_{N\xi} + H_\xi, \quad (15)$$

as

$$\psi_D = \left(1 - \frac{1}{E + i\eta - (T_N + T_P + V_{N\xi} + H_\xi)}\right.$$

$$\times (V_{N\xi} - V_{NP})\Big)\psi_0. \quad (16)$$

In the usual calculation of stripping reaction, approximations are made as $\Psi \to \psi_D$ (Born approximation) and $V_{NP} \to V_{N\xi}$ in Eq. (14).

Exactly the same argument can be made along the line of the above formulation (10) and (11). K-matrix will be written as

$$K_{ba} = (\psi_b, V_{P\xi} \psi_D)$$

$$+ \left(\psi_b, V_{P\xi}\frac{1}{E + i\eta - H}(V_{N\xi} + V_{P\xi})\psi_D\right)$$

$$\approx (\psi_b, V_{P\xi} \psi_D)$$

$$+ \left(\psi_b, V_{P\xi}\frac{1}{E + i\eta - H_a}(V_{N\xi} + V_{P\xi})\psi_D\right)$$

$$\approx (\psi_b, V_{P\xi} \psi_D)$$

$$+ \left(\psi_b, V_{P\xi}\frac{1}{E + i\eta - H_{ai}}(V_{N\xi} + V_{P\xi})\psi_D\right),$$

$$(17)$$

the first approximation being Born approximation corresponding to $\Psi \to \psi_D$ in Eq. (14) and the second one being that corresponding to $V_{NP} \to V_{N\xi}$. There

$$H_a = T_N + T_P + V_{NP} + H_\xi \quad (18)$$

and

$$H_{ai} = T_N + T_P + V_{N\xi} + H_\xi. \quad (19)$$

Thus

Total Errors

$$= \left(\psi_b, V_{P\xi}\frac{1}{E_a + i\eta - H}(E_{ai} - E_a + H - H_{ai})\right.$$

$$\times \frac{1}{E_{ai} + i\eta - H_{ai}}(V_{N\xi} + V_{P\xi})\psi_D\Big)$$

$$= \left(\psi_b, V_{P\xi}\frac{1}{E_a - H}\left\{\left[(V_{NP} - V_N),\right.\right.\right.$$

$$\frac{1}{E_{ai} + i\eta - (T_N + T_P + V_N + H_\xi)}(V_{N\xi} + V_{P\xi})\right]$$

$$+ (V_{N\xi} + V_{P\xi})\frac{1}{E_{ai} + i\eta - (T_N + T_P + V_N + H_\xi)}$$

$$(V_{N\xi} + V_{P\xi})\Big\}\psi_D\Big). \quad (20)$$

In this approximation we see only a part of entrance channel errors in (11) remains.

Since errors can be separated automatically and Φ_a and Φ_b in (11) are easy to take in usual case,

error estimation, at least approximate estimation, can be made easily and vividly in general case also by use of the formula (11).

1) B. A. Lippmann and J. Schwinger, Phys. Rev. **79** (1950), 469.
2) Y. Fujimoto, S. Hayakawa and K. Nishijima, Prog. Theor. Phys. **10** (1953), 113.
3) G. F. Chew and M. L. Goldberger, Phys. Rev. **87** (1952), 778.
4) E. Gerjuoy, Phys. Rev. **91** (1953), 645.
5) Y. Fujimoto, Report of International Conference on Theoretical Physics, held in Koyto, September, 1953.

Note on Meson-nucleon Interaction

Shigeo Minami

Department of Physics, Osaka University

January 31, 1954

Some experimental results with regard to meson-nucleon scattering and photomeson production have shown that the state of angular momentum $J=3/2$ and iostpic spin $I=3/2$ plays the most important role. Here we want to get some insight into the problem of angular momentum, only.

In order to produce the state of $J=3/2$ in the frame of pseudoscalar meson theory, meson-nucleon interaction must be conveyed through p- or d-wave since a nucleon has spin 1/2. Most studies hitherto put forward have dealt with p-wave. It is the aim of this note to examine consequences of the assumption that the interaction responsible for the resonance is d-waves.

If we adopt the isobar model with spin 3/2 as used previously,[1] the parity conservation law leads to the following results: meson and nucleon interact (1) through p-waves in the case of pseudospinor isobar, and (2) through d-waves in the case of spinor isobar.

Composition of angular momenta: nucleon spin $1/2$+meson orbital p or d, gives rise to the two sets of two states with the total angular momentum 1/2, 3/2, 3/2, 5/2, which will be denoted by M, L, L' and K respectively. Nucleon spin functions which have the z-component 1/2 and $-1/2$ are represented

by α and β respectively. Making use of Clebsch-Gordan's coefficients the following relations are easily obtained.

(a) in the case of p-wave interaction

$$
\begin{cases}
L_{3/2}=\alpha Y_{1,1} \\
L_{1/2}=(1/\sqrt{3})[\sqrt{2}\,\alpha Y_{1,0}+\beta Y_{1,1}] \\
L_{-1/2}=(1/\sqrt{3})[\alpha Y_{1,-1}+\sqrt{2}\beta Y_{1,0}] \\
L_{-3/2}=\beta Y_{1,-1}
\end{cases}
\tag{1}
$$

and

$$
\begin{cases}
\alpha Y_{1,0}=(1/\sqrt{3})[\sqrt{2}\,L_{1/2}-M_{1/2}] \\
\beta Y_{1,0}=(1/\sqrt{3})[\sqrt{2}\,L_{-1/2}+M_{-1/2}]
\end{cases}
\tag{2}
$$

(b) in the case of d-wave interaction

$$
\begin{cases}
L'_{3/2}=(1/\sqrt{5})[-\alpha Y_{2,1}+2\beta Y_{2,2}] \\
L'_{1/2}=(1/\sqrt{5})[-\sqrt{2}\,\alpha Y_{2,0}+\sqrt{3}\,\beta Y_{2,1}] \\
L'_{-1/2}=(1/\sqrt{5})[-\sqrt{3}\,\alpha Y_{2,-1}+\sqrt{2}\,\beta Y_{2,0}] \\
L'_{-3/2}=(1/\sqrt{5})[-2\alpha Y_{2,-2}+\beta Y_{2,-1}]
\end{cases}
\tag{3}
$$

and

$$
\begin{cases}
\alpha Y_{2,0}=(1/\sqrt{5})[\sqrt{3}\,K_{1/2}-\sqrt{2}\,L'_{1/2}] \\
\beta Y_{2,0}=(1/\sqrt{5})[\sqrt{3}\,K_{-1/2}+\sqrt{2}\,L'_{-1/2}].
\end{cases}
\tag{4}
$$

On the basis of these relations we can examine the angular distributions for $P+\pi^+\to P+\pi^+$, $\gamma-\pi^0$ processes as well as the total cross section for the former. If we lay the z-axis along the direction of the incident meson, the z-component of the total angular momentum is 1/2 or $-1/2$ corresponding to the initial state $\alpha Y_{1,0}$ or $\beta Y_{1,0}$ in case (a) and $\alpha Y_{2,0}$ or $\beta Y_{2,0}$ in case (b) respectively. If the process which involves the isobar state with spin 3/2 resonates and the $L(L')$-state only predominates, then the scattered waves are described by $L_{1/2}$ or $L_{-1/2}$ ($L'_{1/2}$ or $L'_{-1/2}$). The resulting angular distribution are proportional to

$$2|Y_{1,0}|^2+|Y_{1,1}|^2 \qquad \text{in case (a)},$$

$$2|Y_{2,0}|^2+3|Y_{2,1}|^2 \qquad \text{in case (b)}.$$

It is remarkable that there results the same angular distribution $(3\cos^2\theta+1)$ in both cases, the $\cos^4\theta$ and related terms in case of (b) cancelling out.

Let us compare the magnitude of the total cross section in case (a) with that in case (b) taking into account the resonance scattering only. In the well known formula

$$\sigma_T = (4\pi\, \Lambda^2)\sum_l (2l+1)\sin^2\delta_l$$

we put $\delta_l = 90°$ for $l=1$ or $l=2$ and neglect other terms, so that

$$\sigma_T = C_1(4\pi/\Lambda^2)\times 3 \qquad \text{in case (a)},$$

$$\sigma_T = C_2(4\pi/\Lambda^2)\times 5 \qquad \text{in case (b)}$$

where $C_i (i=1, 2)$ are the probabilities with which the $L(L')$-state appears in the initial state. From eqs. (2) and (4) we get $C_1 = 2/3$ and $C_2 = 2/5$. Therefore the magnitudes of total cross section in both cases are also the same.

By similar considerations, the angular distributions for $\gamma - \pi^0$ process are found to be proportional to

$$3|Y_{1,1}|^2 + (1/3)2|Y_{1,0}|^2 + |Y_{1,-1}|^2] \qquad \text{in case (a)},$$

$$|Y_{2,1}|^2 + 4|Y_{2,2}|^2 + (1/3)[3|Y_{2,-1}|^2 + 2|Y_{2,0}|^2] \qquad \text{in case (b)}$$

which are of the same form $(2 + \sin^2\theta)$ in both cases.

Thus we cannot decide whether the p or the d wave resonance actually occurs, as long as we pay attention only to meson-nucleon scattering and neutral meson production by γ-ray. We must look for a process on which some distinctions can be made between them. Watson and Brueckner[2] had investigated the problem of meson production in nucleon-nucleon collision from the phenomenological point of view and showed that the experimental facts, especially the small cross section for $P+P\to P+P+\pi^0$, could be explained satisfactorily if meson were emitted in p-wave. At that time beam energy was of the order of 340 Mev and the cross section for (PP, π^0) was very small; but the subsequent experiments[3] at higher energies (430 Mev) have shown that it increases rapidly. As beam energy increases, the $L(L')$-state becomes more important for this reaction.

In such an energy range, let us observe the energy spectrum of emitted meson by this process and pay attention to the region which lies below and in the immediate neighbourhood (about 10 Mev) of the maximum energy. So far as we are concerned with this region of energy spectrum, it may be expected that the two final nucleons are in $^1S_0$ state since emitted meson carries away the most part of incident energy.

(a) In this case it is a good approximation to regard the meson as emitted in the center of mass system with relative momentum $l=1$.

As pointed out by Watson and Brueckner, this reaction is forbidden by Pauli's exclusion principle, parity and angular momentum conservation. As a result, in this energy region the cross section turns out to be very small and to have such a form as (a)-curve in Fig. 1.

(b) When L' consists of nucleon spin 1/2 and d-meson, the transition :

$$^3P_2 \longrightarrow {}^1S_0 + (d\text{-wave meson})$$
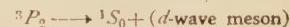
is allowed. Thus the cross section in the region under consideration will have such a form as (b) in Fig. 1.
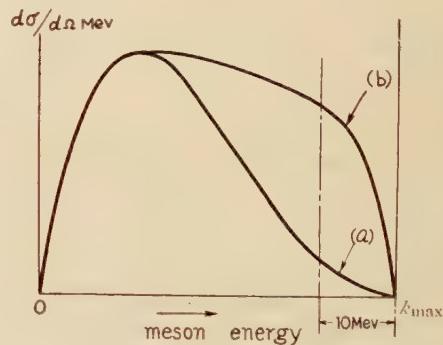
Fig. 1

Thus we will be able to expect the possibility of discrimination between (a) and (b). It may be surmised that the pseudovector coupling theory will lead to similar circumstances as case (a). The whole view of energy spectrum will be worked out in near future.

In conclusions, the author should like to express his thanks to Prof. K. Husimi, Prof. Z. Koba, Prof. S. Hayakawa and the members of the research groups for the theory of elementary particles in Osaka for their valuable discussions.

1) S. Minami, T. Nakano, K. Nishijima, H. Okonogi and E. Yamada, Prog. Theor. Phys. **8** (1952), 531.
2) K. M. Watson and K. A. Brueckner, Phys. Rev. **83** (1951), 1.
3) Marshall, Marshall, Nedgel, Warshaw, Phys. Rev. **88** (1952), 3, 632,

Mass Spectrum of Elementary Particles I
—— *Eigenvalue Problem in Space-time* ——

Hiroshi ENATSU*

Department of Physics, Columbia University
New York, New York, U.S.A.

(Received December 7, 1953)

In order to avoid the divergent self-energies of quantum field theory and at the same time to get the mass spectrum of elementary particles, a new method is discussed.

A generalized wave equation which contains a function of a time-like parameter and four variables in the space-like domain, is assumed. In the wave equation the self-energy of a particle is considered to be a self-potential which is dependent only of the space-like variables. An investigation is made of how the eigenvalue problem for the masses of elementary particles is set up, and is solved by assuming a simple potential form. It will be shown that a discrete mass spectrum is obtained for a negative self-potential under some approximations.

§ 1.   Introduction

It is a remarkable fact that many new particles have been found not only in the recent experiments of cosmic rays,[1] but also in the research of the artificial production of heavy particles by the Brookhaven cosmotron. It is felt, in the face of the new experimental development, that it would be necessary to examine more fully the mass problem of elementary particles from general viewpoint in the framework of the present field theories. This problem has been already discussed by many people.[2a,b,5] However, whether such theories can be made to give the observed mass spectrum of particles and at the same time to remove divergent self-energies is still a question that must be settled by further investigations. Therefore, it may be worth while to carry out this program at the present stage.

The recent developments of quantum electrodynamics have made it possible to treat many problems in a completely covariant way. Nevertheless, since the mass of a particle has been considered to be a parameter at all stages, the use of the so-called covariant theory seems to be inconvenient for our purpose of investigating the problem of the mass spectrum. The result of recent experiments indicates that the mass of the elementary particle has to be quantized in some ways which are not known yet. If this is the case, it may be necessary to set up a new mathematical formalism for the eigenvalue problem of mass. As to the mass of the elementary particle itself, we have another trouble. As is well known, so far the divergent mass corrections on account of interaction have been

---

* Now at Kyoto University, Kyoto, Japan.

disposed of by means of the renormalization technique[3] and the compensation method.[4] However, these procedures seem to be of phenomenological nature. For example, in the latter case one can get a relation between the masses of elementary particles involved, but one cannot know each mass value separately. At any rate, as part of the experimental mass, the divergent mass correction must be made finite. In this connection, the so-called non-local interaction theory[5] and the non-local field theory have been suggested by several authors.[2a] These theories have, however, met with considerable difficulties when some conditions, such as relativistic invariance, causality and convergence, have been examined. It seems that the unsatisfactory features in these theories are partly due to the formal and mathematical way of approach. In fact, there seems to be too little physical approach at present to look for the stumbling-block in the present field theory. It is clear why the viewpoint of non-localizability has been adopted so far in the recent investigations. In attacking divergence problems, one would expect that the point models for the elementary particle could be modified in such a way that one could get convergence by introducing form factors. Consequently, the theories contain arbitrary functions which would not be determined by any physical reason. Furthermore, the step of investigation in these theories is divided into two parts. Namely, one is the problem of convergence, and the other is the problem of mass spectrum. In most of the works that have been done, they have confined their attention to the problem of convergence so that it has not been clear whether the so-called convergence factors would give rise to the reasonable mass spectrum of elementary particles. It is to be stressed that, since the self-energy appears as part of mass, any attempt to get rid of the infinity must be directly connected to the eigenvalue problem of mass. In order to see how this can be done, we need a new treatment which may be different from that of the usual non-local theory, but which may include the idea of the non-localizability somehow.

This paper will describe a new formulation for this problem, in which the mass spectrum and self-energy will be considered simultaneously in a relativistic way. The starting point of our analysis is an extended form of the ordinary wave equation. It seems that there are two ways in order to generalize the Schrödinger equation in the sense of relativistic quantum mechanics. One is the so-called Tomonaga, Schwinger and Feynman formalism. As for this extension there is no need to go further because of the reason mentioned above. The other, a better alternative for us, is a proper time formulation[6] which introduces some invariant parameters instead of time t. With this formulation many problems, such as the problems of gauge invariance and vacuum polarization, have been treated. It is unsatisfactory to find that there still remains the above difficulty. However, in spite of this defect, if one considers more fully the physical significance of this formulation, one sees that there seems to be a possibility of avoiding the trouble and of finding out a way to get a solution for the mass spectrum.

First, we will assume a generalized Schrödinger equation in space-time for one-body problem. This can be done in a completely relativistic way. As one will see later, the development of a system with an invariant parameter is governed by a generalized Hamiltonian which may be assumed to be a mass operator in space-time. With the

generalized Hamiltonian the equations of motion are invariant with respect to Lorentz transformations. In this connection, the definition of a wave function will also be changed. An important question which has been presented from the earlier stage of the development of quantum mechanics, is whether it is necessary to define completely relativistic probablity amplitudes. Dirac[7] proposed such a concept of states in order to extend the ordinary state vector in space-time. However, within the frame of the quantum mechanics, it was of no use to do so. On the contrary, we shall see later that in the case of the mass-spectrum problem, we need a formulation which includes the extended state vectors in the sense of Dirac, and which is a function of an invariant parameter. This is connected intimately with the following facts. In the ordinary eigenvalue problem the energy of a system, which is a component of an energy-momentum vector in space-time, is specified, so that the concept of stationary states has been associated with a state vector which is a special function of time. On the other hand, in the case of the mass problem, the eigenvalue must be a scalar in space-time. Accordingly, the state has to depend particularly not on the time but on the invariant parameter. Secondly, as is well known, in the case of a Dirac electron, the space components of angular momentum tensor have been merely taken into account, while the other components have been discarded. Therefore, in order to be consistent with the above extension of the state vector, it is necessary to put the usual angular momentum tensor into its complete form in space-time. Thirdly, as to the treatment of the divergent self-energies, a new method will be introduced. Namely, it will be assumed that the divergent self-energy will be regarded as a kind of self-potentials[8] in terms of which an eigenvalue problem for the mass will be solved in space-time.

We can express our idea in the following way. It is well known that the divergent self-energy appearing in quantum field theory may reveal certainly the apparent lack of consistency. Hence, one may expect that in some future theory the divergent terms will be made finite. However, we think that such a difficulty may be looked upon merely as presenting an insufficient treatment of the problem connected with the masses of elementary particles. As far as we ristrict ourselves to the ordinary consideration of mass, we can not avoid the divergence difficulty. However, by taking the space-time viwpoint for it, we may rather make use of the divergence to get the mass spectrum of elementary particles.

As a particular illustration, we will show how the mass spectrum of fermions may be treated in the space-time formulation by assuming a rest-mass and a simple potential which arises from the interaction with scalar photons. The general idea of this work may be readily extended to the other problems of divergence difficulties.

§ 2. The generalized wave equations in space-time

We begin by considering the generalization of the Schrödinger equation for one-body problem. The Schrödinger equation for the $\psi(x, y, z, t)$-vector representing the state for a system has the form

$$i\frac{\partial \psi(x,\ y,\ z,\ t)}{\partial t}=H\psi(x,\ y,\ z,\ t),\qquad \hbar=c=1. \tag{1}$$

This must be extended to take into account the mass of an elementary particle, which involves a rest (or proper)-mass and the mass correction on account of interaction, as an eigenvalue of an eigenvalue problem in space-time. Of course, according to Tomonaga[9] and Schwinger[10], the non-relativistic wave equation can be brought into a relativistic invariant form :

$$i\frac{\partial \Psi(\sigma)}{\partial \sigma(x)}=H_I(x)\Psi(\sigma). \tag{2}$$

However, in this theory emphasis has been placed on the energy operator, while the mass has been regarded merely as a parameter.

In order to get a wave equation which is adequate to our purpose, let us start with classical relations

$$(p_\mu\,p_\mu)=p_1^2+p_2^2+p_3^2-E^2=-m^2, \tag{3}$$

and

$$(x_\mu x_\mu)=x_1^2+x_2^2+x_3^2-t^2=+s^2. \tag{4}$$

These relations suggest another way of generalizing the Schrödinger equation (1). That is to say, instead of taking the energy and the time as a Hamiltonian and a variable of differentiation, we define an invariant Hamiltonian M which means a mass operator, and introduce two parameters τ and σ which correspond to the classical parameters $-s^2$ and $+s^2$ respectively. The real parameters τ and σ specify respectively the development and behavior of a relativistic wave function $\Psi(x_\mu(\sigma),\ \tau)$ in time-like and space-like regions. It is also assumed that in a complex plane τ runs along a real axis while the range of σ lies on an imaginary axis. It should be noted that two parameters are defined in the invariant time-like and space-like domains in the theory of relativity.

We assume a generalized wave equation as follows :

$$i\frac{\partial \Psi(x_\mu(\sigma),\ \tau)}{\partial \tau}=M(x_\mu(\sigma))\Psi(x_\mu(\sigma),\ \tau). \tag{5}$$

In this equation, $\Psi(x_\mu(\sigma),\ \tau)$ means that the wave function depends on the space-like variables $x_\mu(\sigma)$ and the time-like parameter τ whose significance will be given later. We suppose a state of the system can be specified by these five variables. We will look upon $M(x_\mu(\sigma))$ as an operator which depends merely on the space-like variables $x_\mu(\sigma)$. Correspondingly, in the Heisenberg representation we have

$$\frac{dx_\mu}{d\tau}=-i\,[x_\mu,\ M], \tag{6}$$

and

$$\frac{dp_\mu}{d\tau}=-i\,[p_\mu,\ M], \tag{7}$$

where

$$[x_\mu,\ p_\nu]=i\,\delta_{\mu\nu}. \tag{8}$$

It is to be noted that the relation (8) shows the commutation rule with respect to the four-dimensional operators λ_μ and p_ν, which has now a perfect space-time form. We shall solve the equation (5) in a manner analogous to that of the ordinary quantum mechanics. The most obvious way of doing this is by separating the wave functions. In the next section we will define a self-potential which is independent of the parameter τ, it is then possible to express the solution of (5) as a product of functions of $x_\mu(\sigma)$ and τ separately. It is therefore reasonable to write generally

$$\Psi(x_\mu(\sigma),\ \tau) = \exp(im'\tau)\psi(x_\mu(\sigma)), \tag{9}$$

where $\psi(x_\mu(\sigma))$ shows the wave function which is a function of the space-like variables $x_\mu(\sigma)$ only, and l is one for a Dirac particle or two for a Klein-Gordon particle. The procedure of separation such as (9) is clearly relativistic, because it corresponds to the division of the whole of space-time into two invariant domains. The expressions (5) and (9) are the starting points of our consideration. To make clear the physical idea let us consider a simple case that we have a particle in space-time. Usually, in quantum mechanics a wave function is specified by a flat space-like surface $t=$ constant. The energy operator H is an infinitesimal operator in the direction of time t. Thus the so-called Hamiltonian formalism is plainly non-relativistic. We can achieve the desired covariance by considering a space-like surface which is defined by a condition $\tau=$ constant, and by introducing the generalized operator M. Of course, if necessary, we may define a point $x_\mu(\tau)$ within the light cone whose vertex is at the origin* as in the cases of (6), (7) and (8). However for our purpose the only thing we have to do within the light cone is not to localize a point $x_\mu(\tau)$ but to specify a space-like surface. In the case of the region outside the light cone, it is to be expected that, since the operator M may contain the components of some operators, the wave function $\psi(x_\mu(\sigma))$ may involve space-like coordinates $x_\mu(\sigma)$ explicitly. As is easily seen, the separation (9) is a natural extension of that in quantum mechanics, which is of the form

$$\Psi(r,\ t) = \exp(-iEt)\varphi(r). \tag{10}$$

Hence, the function $\psi(x_\mu(\sigma))$ will be connected with the probability amplitude of the particle outside the light cone.

Now it may be useful, as we will see later, to derive relations between the parameter τ and 4-dimensional classical distances by means of the equations (6) and (7). Generally, we require that the relation

$$\frac{dF}{d\tau} = -i[F,\ M] \tag{11}$$

be valid for every physical quantity F which dose not explicity depend on the parameter τ. The generalized Hamiltonians for the free Dirac and Klein-Gordon particles are respectively

---

* As will be shown in the relation (49), we suppose one of the points $x_\mu(\tau)$ and $x_\nu(0)$ to be zero, i.e. $x_\mu(0) = 0$.

$$M = \gamma_\mu p_\mu, \qquad \gamma_\mu \gamma_\nu + \gamma_\nu \gamma_\mu = -2\delta_\mu \ , \tag{12}$$

and

$$M = p_\mu^2. \tag{13}$$

In the case of the Dirac particle we now apply the relation $| \ | \ | \ | \ |^{\cdots}$, p_μ, x_μ and γ_μ. The results are

$$\frac{dM}{d\tau} = 0, \qquad M = \text{const.} = -m, \tag{14}$$

$$\frac{dp_\mu}{d\tau} = 0, \qquad p_\mu = \text{const.}, \tag{15}$$

$$\frac{dx_\mu}{d\tau} = \gamma_\mu, \tag{16}$$

and

$$\frac{d\gamma_\mu}{d\tau} = 2i\gamma_\mu m + 2ip_\mu. \tag{17}$$

By differentiating with respect to τ, we get from (17),

$$-i \frac{d^2 \gamma_\mu}{d\tau^2} = 2 \frac{d\gamma_\mu}{d\tau} m. \tag{18}$$

Integrating this expression we have

$$\frac{d\gamma_\mu(\tau)}{d\tau} = \left(\frac{d\gamma_\mu}{d\tau}\right)_{\tau=0} e^{2im\tau}. \tag{19}$$

From (16) (17) and (19) the following integration can be obtained,

$$(x_\mu(\tau) - x_\mu(0)) = -\frac{1}{4}\left(\frac{d\gamma_\mu}{d\tau}\right)_{\tau=0} e^{2im\tau} \frac{1}{m^2} - \frac{p_\mu \tau}{m} \ . \tag{20}$$

The first term on the right-hand side shows the effect corresponding to the "Zitterbewegung". In this paper we will not consider the exact form of the 4-dimensional distance. Then, if all small terms in (20) representing quantum effects are neglected to get an approximate expression for the space-time distance, the last form can be written as

$$(x_\mu(\tau) - x_\mu(0))^2 \sim -\tau^2. \tag{21}$$

This relation reveals that in this approximation the two different parameters $-s^2$ in (4) and $-\tau^2$ in (21) have the same relation to the space-time distance. Since the parameter τ is real, one easily sees the time-like character of the distance. The same method can be applied to the Klein-Gordon particle for which the Hamiltonian is given by the relation (13). Then we see that the corresponding expression will be

$$(x_\mu(\tau_1) - x_\mu(0))^2 \sim -4m^2 \tau_1^2. \tag{22}$$

So far we have limited ourselves to the cases of free particles. It is possible that a detailed analysis could show that the relations (21) and (22) are valid for the particles

in an external potential. Thus we have found a special form of the space-time distance, which would be used only in the time-like region. However, in the previous section we have introduced also another parameter σ in space-time, and the relation between the σ and the space-time distance has not been analyzed. At first sight there seems to be no ground for reasoning such a relation. However, there is even reason to believe that following expressions could be assumed,

$$(x_\mu(\sigma) - x_\mu(0))^2 \sim \sigma^2, \tag{23}$$

and

$$(x_\mu(\sigma_1) - x_\mu(0))^2 \sim 4m^2\sigma_1^2. \tag{24}$$

That this possibility exists can be seen from the classical relation (4) which shows the complete symmetry with respect to the two parameters.

§ 3. Self-potentials

We now proceed to consider the wave equation for a system composed of a particle and its self-fields. This we do in two steps. First we wish to consider the meaning of the so-called mass type self-energy in space-time representations. Next we shall set up the wave equation for the mass eigenvalue problem.

As is well known, in the momentum representation the self-energy of the Dirac particle caused by its electromagnetic field is given[4] by

$$\delta m = m \frac{3}{2\pi} \left(\frac{e^2}{4\pi\hbar c} \right) \log \frac{K + K_0}{M} + \text{finite terms}, \tag{25}$$

where $K_0' = (K^2 + M^2)$ and $K \to \infty$.

The important feature of this expression is that there appears divergence at large momenta due to the insufficiently rapid decrease of the integrand in a momentum integral, and physically the difficulty is ascribed to the large momenta of photons involved in intermediate states. In many works that have been done on the problem of self-energy, the stress has been laid on its momentum representation. It becomes convenient at this stage to use the space-time representation of self-energy : the representation which is expressed in terms of the parameters τ or σ. We may then ask a fundamental question : How does the difficulty arise in space-time ? We shall discuss this problem by considering a simple example of the Dirac particle interacting with scalar photons. The reason of this choice is that we have a negative self-energy as a result of interactions, which means an attractive self-action in space-time. It is most convenient to write the self-energy by making use of the formulation proposed by Schwinger.[6] According to his method the Dirac equation is

$$-i\gamma_\mu \partial_\mu \psi(x) + \int \mathfrak{M}(x, x') \psi(x') (dx') = 0, \tag{26}$$

where

$$\mathfrak{M}(x, x') = m_0 \delta(x - x') + ig^2 G(x, x') D_+(x, x'), \tag{27}$$

and m_0 is the rest-mass, and g is the coupling constant. Here the functions $G(x, x')$ and $D_+(x, x')$, which are Green's functions for the Dirac particle and scalar photon respectively, are written as

$$G(x, x') = \frac{1}{(4\pi)^2} \int_0^\infty \tau_1^{-2} \exp(-im^2\tau_1) \left[-\frac{\gamma_\mu(x_\mu - x'_\mu)}{2\tau_1} + m \right] \exp\left\{ i \frac{(x_\mu - x'_\mu)^2}{4\tau_1} \right\} d\tau_1$$

(28)

and

$$D_+(x, x') = \frac{1}{(4\pi)^2} \int_0^\infty \rho_1^{-2} \exp\left\{ i \frac{(x_\mu - x'_\mu)^2}{4\rho_1} \right\} d\rho_1.$$

(29)

It must be noted that these Green's functions are derived by means of the Hamiltonian (13). Hence, the relation (22) is assumed for the parameters τ_1 and ρ_1 involved. The procedure here described for setting up a wave equation involving infinite self-energies is very similar to that of the non-local interaction theory. In fact, it is evident that the equation (26) is of the so-called non-local type.[5] However, the situation of the non-locality · is slightly different for the cases of the usual non-local theories. Coventionally, we may classify the non-local theories into three types as to the order of the coupling constants. First of all, we have the non-local field theory in which the wave function $\psi(x)$ itself is non-localizable in the sense of Yukawa.[20] Thus we can call this as the first kind non-local theory. As the second kind of types one might take the non-local interaction and non-local action theories.[5] It may be pointed out that the occurrence of non-locality is in the interaction term which is linearly proportional to coupling constants. As we know, in this case the form factor should be taken into account to get convergence. In contrast to the cases mentioned above one finds that the equation (26) contains a non-local term which begins with the second-order term in the coupling constant. Now it is easy to see that first-order interaction effects have no direct physical significance. This fact seems to give us the reason why it is very difficult to find the physical reason of form factors in the case of the second kind non-local theory. However, in our case the wave function itself will play the role of the form function which will suppress the infinity of the self-energy at the light cone.

To complete the picture of this alternative procedure for dealing with the non-local terms, we shall carry out the modification of the equation (26). After some calculations we obtain

$$(-i\gamma_\mu\partial_\mu + m)\psi(x) = 0,$$

(30)

where

$$m = m_0 + \delta m,$$

(31)

$$\delta m = -\frac{g^2}{4\pi}\left(\frac{m}{4\pi}\right)\int_0^\infty \tau_1^{-1}d\tau_1 \int_0^{\tau_1} \tau_1^{-1}\left(1 + \frac{w}{\tau_1}\right)\exp\{-im^2(\tau_1 - w)\}\,dw,$$

(32)

and

$$\frac{1}{w} = \frac{1}{\tau_1} + \frac{1}{\rho_1}.$$

(33)

Thus we have expressed the equation (30) in terms o f τ_1 and w which are time-like parameters. On the other hand, in the previous section, it was shown that from the equations (5) and (9) the wave function $\psi(x_\mu(\sigma))$ was the space-like part of the wave function $\Psi(x_\mu(\sigma), \tau)$. Therefore in the expression (30) the wave function $\psi(x)$ should be of this type. However it has not been formulated in such a way that it is certainly obvious that the $\psi(x)$ in (30) just corresponds to that in the equation (9). At this point a considerable reinterpretation of the equation (30) would be needed. For that purpose the following transformations due to Schwinger[6] will be convenient;

$$1 - \frac{w}{\tau_1} = u,\tag{34}$$

and

$$\tau_1 \rightarrow -i\sigma_1.\tag{35}$$

Making use of these transformations, from (32) one gets the expression for δm :

$$\delta m = -\frac{g^2}{4\pi}\left(\frac{3}{8\pi}\right)m\int_0^\infty \frac{e^{-m^2\sigma_1}}{\sigma_1}d\sigma_1 - \frac{g^2}{4\pi}\left(\frac{7}{16\pi}\right)m.\tag{36}$$

We can understand the meaning of the variable changes (34) and (35) in the following way. We suppose that in the mass correction δm the path of integration of the time-like parameter τ_1 is along the real axis with a pole at the origin. One may regard δm as a function of the variables on the complex plane (τ_1, σ_1). In order to show δm explicity as real quantity, the path of integration may be deformed. That is to say, for the parameter τ_1 we take the path,

$$\tau_1: \qquad \epsilon \rightarrow \infty, \qquad (\epsilon \sim 0),\tag{37}$$

on the real axis. The next contour is part of a circle in a clock-wise direction,

$$\theta: \quad 0 \rightarrow -\frac{\pi}{2},\tag{38}$$

until it lies along the imaginary axis which corresponds to the range of the space-like parameter σ_1,

$$\sigma_1: \quad -\infty \rightarrow -\epsilon.\tag{39}$$

There is no contribution from the path (38). However, there remains an additive constant in the integration on the path of a small quadrant with the radius ϵ around the pole, which will be neglected in the following discussion. One thus obtains the mass correction δm as an integral over the space-like parameter σ_1 with the result (36). In other words, we have found the space-like representation of the self-energy. As mentioned above, the space-time point x_μ in the equation (30) may be assumed to be outside the light cone. Thus we are now considering an eigenvalue problem of mass in which the self-energy might play the role of self-potential in this domain of space-time. However, in the expression (36) a logarithmic divergence arises when we approach the light cone in the space-like direction ($\sigma_1 \rightarrow 0$). We must therefore study how to avoid such a

difficulty. It seems that the situation is analogous to that in the case of elementary quantum mechanics, where the Coulomb potential becomes infinite at the origin in the Dirac equation. However, as we know, we have finite energy levels for the hydrogen atom. The essential point to be noted is that the singularity of the Coulomb potential at the origin is avoided by the fact that the position probablity density of the electron at the origin is zero. Therefore in the present treatment it is an interesting trial to reinstate the idea of the relation between the potential and wave function. For this purpose there seems to be an adequate expression of the Coulomb potential in the 3-dimensional space. We recall the representation of it as a result of the integration of field strength from infinity to the point r considered,

$$-\frac{e^2}{r} = e^2 \int_\infty^r \frac{1}{r^2}\,dr = -e^2 \int_r^\infty \frac{1}{r^2}\,dr. \tag{40}$$

This relation suggests that, even in our case instead of taking the limiting value $\sigma_1 = 0$ in the expression (36), the following self-potential may be adopted in the equation (30)

$$V(\sigma_1) = -\frac{g^2}{4\pi}\left(\frac{3}{8\pi}\right)m \int_{\sigma_1}^\infty \frac{e^{-m^2\sigma_1}}{\sigma_1}\,d\sigma_1. \tag{41}$$

In other words, we assume that $V(\sigma_1)$ is a function of the variable σ_1 (self-action at a distance) and is inserted in the wave equation whose wave function is $\psi(x_\mu(\sigma))$,

$$(\gamma_\mu \partial_\mu + \mu_0 + V(\sigma_1))\psi(x_\mu(\sigma)) = 0, \tag{42}$$

with

$$\gamma_\mu \gamma_\nu + \gamma_\nu \gamma_\mu = 2\delta_{\mu\nu}, \tag{43}$$

$$\sigma = 2m\sigma_1, \tag{44}$$

and

$$\mu_0 = m_0 - \frac{7g^2}{64\pi^2}m. \tag{45}$$

The relation (44) which shows the change in gauge arises from the different definitions of space-time distances for the potential (41) and the equation (42).

It should be remarked that the definition of the potential $V(\sigma_1)$ is the analog of the classical potential, but this does not mean the direct correspondence between the quantum mechanical and classical potentials. This procedure may be looked upon only as that given by the correspondence principle.

Now, let us consider $V(\sigma_1)$ which is a monotonous but complicated function of the parameter σ_1. Briefly it can be expressed in terms of the exponential integral $\mathrm{Ei}(-x)$. We shall be content to write its asymtotic forms as follows:

$$V(\sigma_1) = -\frac{g^2}{4\pi}\left(\frac{3}{8\pi}\right)m I(\sigma_1), \tag{46}$$

where

$$I(\sigma_1) \sim e^{-m^2\sigma_1}\left(\frac{1}{m^2\sigma_1} - \frac{1}{m^4\sigma_1^2} + \frac{2!}{m^6\sigma_1^3} - \cdots\cdots\right), \quad m^2\sigma_1 \gg 1, \tag{47}$$

and

$$I(\sigma_1) \sim -\gamma - \log m^2\sigma_1 + \frac{m^2\sigma_1}{1!} - \frac{m^4\sigma_1^2}{2 \cdot 2!} + \cdots\cdots , m^2\sigma_1 \ll 1, \qquad (48)$$

(γ : Euler constant).

Here one obtains a logarithmically divergent result for the small value of σ_1.

In this way the form of the wave equation, as well as the self-potential, is finally given in the space-time representation. As a final comment of this section, we remark that, as stated above, the wave function $\psi(x_\mu(\sigma))$ may be regarded as the probability amplitude as in the case of elementary quantum mechanics so that our formalism may permit an extended particle model outside the light cone. Therefore the formulation given here on the basis of the current local field theories, may involve the idea of non-localizability of the particle as a result of the interaction with itself in the sense of Yukawa[2a] only in the space-like domain.

§ 4. Angular momenta in space-time

We may now proceed to solve the eigenvalue problem outside the light cone. We shall not be much concerned with the exact solution because of mathematical troubles. In order to simplify further discussion, we shall employ a semiclassical approximation. In discussing the space-time distance, we make a convenient substitution to perform practical calculations. Namely, we suppose the initial values $x_\mu(0)$ to be zero, so that (23) becomes

$$x_\mu^2(\sigma) = \sigma^2. \qquad (49)$$

With this definition we introduce the following polar coordinates[10] outside the light cone ;

$$
\begin{aligned}
x_1 &= R \cosh \omega \sin \theta \cos \varphi, & 0 &< R < \infty, \\
x_2 &= R \cosh \omega \sin \theta \sin \varphi, & -\infty &< \omega < \infty, \\
x_3 &= R \cosh \omega \cos \theta, & 0 &\leq \theta < \pi, \\
x_0 &= R \sinh \omega, & 0 &\leq \varphi < 2\pi.
\end{aligned}
\qquad (50)
$$

Then we have from (49)

$$x_\mu^2 = R^2, \qquad (\text{i.e. } \sigma = R.). \qquad (51)$$

The volume element is given by

$$dV = dx_0 dx_1 dx_2 dx_3 = R^3 \cosh^2\omega \sin \theta \, dR \, d\omega \, d\theta \, d\varphi. \qquad (52)$$

Our next task is to construct the general angular momentum tensor in the new represent-ation. We define the angular momentum operator in space-time as follows :

$$L_{\mu\nu} = x_\mu p_\nu - x_\nu p_\mu, \qquad (\mu, \nu = 1, 2, 3, 4). \qquad (53)$$

The process will be facilitated by introducing two components of $L_{\mu\nu}$, and their squares, which are given by

$$M_a = I_{bc}, \qquad (a,\ b,\ c=1,\ 2,\ 3;\ \text{cyclic}),\quad (a \gtrless b \gtrless c), \tag{54}$$

$$N_a = L_{4a}, \tag{55}$$

$$M^2 = (M_a M_a), \tag{56}$$

$$N^2 = (N_a N_a), \tag{57}$$

and

$$K^2 = \frac{1}{2}(L_{\mu\nu} L_{\mu\nu}) = (L:L). \tag{58}$$

With these definitions, the commutation laws for them are easily deduced,

$$[M_a,\ M_b] = [N_a,\ N_b] = i\, M_c, \tag{59}$$

$$[M_a,\ N_b] = [N_a,\ M_b] = i\, N_c, \tag{60}$$

$$[N_a,\ N_a] = [M_a,\ M_a] = 0, \tag{61}$$

and

$$[K^2,\ M_a] = [K^2,\ N_a] = 0. \tag{62}$$

The method of solving the angular parts of the equation (42) is in many respects similar to that in the case of the Dirac equation for the hydrogen atom. The generalized operators which represent the squares of the angular momenta are found to be

$$K^2 = \frac{\partial^2}{\partial \omega^2} + 2\tanh\omega\frac{\partial}{\partial\omega} + \frac{M^2}{\cosh^2\omega}, \tag{63}$$

where

$$M^2 = -\left(\frac{\partial^2}{\partial\theta^2} + \cot\theta\frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right). \tag{64}$$

We separate the radial and angular parts of the wave function by substituting

$$\psi(x) = \varphi(R)\, W(\omega)\, Y_{lm}(\theta,\ \varphi). \tag{65}$$

As for the wave functions for the angles θ and φ, it is evident that we can take the ordinary expression ;

$$Y_{lm}(\theta,\ \varphi) = N_{lm} P_l^m(\cos\theta)\, \Phi_m(\varphi). \tag{66}$$

The differential equation satisfied by the $W(\omega)$ is

$$K^2 W(\omega) = \lambda W(\omega), \qquad (\lambda:\ \text{eigenvalue}), \tag{67}$$

so that

$$\frac{d^2 W}{d\omega^2} + 2\tanh\omega\frac{dW}{d\omega} + \left\{\frac{l(l+1)}{\cosh^2\omega} - \lambda\right\} W = 0. \tag{68}$$

It is convenient to substitute $B = W\cosh\omega$ for W, then we get

$$\frac{d^2 B}{d\omega^2} + \left\{\frac{l(l+1)}{\cosh^2\omega} - (\lambda+1)\right\} B = 0. \tag{69}$$

The functions $W(\omega)$ and $B(\omega)$ are normalized so that

$$\int_{-\infty}^{\infty} W^* W \cosh^2 \omega \, d\omega = \int_{-\infty}^{\infty} B^* B d\omega = 1. \tag{70}$$

The solutions for the equation (69) are given by[11]

$$B_q^q(\omega) = \left[\frac{2^{1-2q}\Gamma(2q)}{\Gamma(q)\Gamma(q)}\right]^{1/2} \cosh^{-q}\omega, \qquad q > 0, \tag{71}$$

and

$$B_l^q(\omega) = \left[(l-q)(l+q)\right]^{-1/2}\left(l \tanh \omega - \frac{d}{d\omega}\right) B_{l-1}^q(\omega). \tag{72}$$

The corresponding eigenvalues are easily seen to be

$$\lambda = (q-1)(q+1), \tag{73}$$

$$l - q = 0, \ 1, \ 2, \cdots, \qquad q \leqq l. \tag{74}$$

For $q < 0$ the solutions are not independent of those for $q > 0$. In the case of $q=0$, however, we have no eigenfunctions which satisfy the condition (70). Thus q can assume the positive integer values. In this connection it is important to note with the solution for the angle ω that the solutions which are physically acceptable are obtained only when we limit ourselves to the domain outside the light cone. In fact, we might consider such an eigenvalue problem within the light cone, but it would turn out that there would be no solution at all, which satisfies the quadratically integrable condition. Hence it is now clear why we have assumed the self-potential to be a function of the space-like parameter σ_1.

As is well known, the above argument is incomplete for the case of the Dirac particle. We must take into account the spin operator $\sigma_{\mu\nu}$. We therefore start by defining the total angular momentum which is a constant of motion,

$$J_{\mu\nu} = L_{\mu\nu} + \frac{1}{2}\sigma_{\mu\nu}, \qquad \frac{dJ_{\mu\nu}}{d\tau} = 0. \tag{75}$$

From equations (58)-(62) one gets

$$(\sigma : L)(\sigma : L) = (L : L) - 2(\sigma : L). \tag{76}$$

Making use of this relation, we obtain

$$\left(L + \frac{1}{2}\sigma\right)^2 + \frac{3}{4} = \left[(\sigma : L) + \frac{3}{2}\right]^2. \tag{77}$$

Finally, one gets

$$T^2 = J^2 + \frac{3}{4}, \tag{78}$$

with

$$T = \left[(\sigma : L) + \frac{3}{2}\right]. \tag{79}$$

However, the operator T is not commutable with the matrix $\gamma_\mu p_\mu$ in the Hamiltonian, and so is not a constant of motion. It is easy to find a commutable operator. That is to say, we may now employ an operator

$$Q_0 = \gamma_5 [(\sigma : L) + 3/2],\tag{80}$$

which satisfies the relations

$$T^2 = Q_0^2,\tag{81}$$

and

$$[\gamma_\mu p_\mu, Q_0] = 0.\tag{82}$$

The last relation reveals that the Q_0 is a constant of motion. The eigenvalue Q of the operator Q_0 can be inferred in the following way. With the help of equations (78) and (81) one sees that

$$Q_0^2 = J^2 + 3/4.\tag{83}$$

Referring to the condition (73), one can consider that the operator J^2 has the eigenvalue

$$(j-1)(j+1) + 1/4,\tag{84}$$

where the term $1/4$ is introduced for convenience (cf. Ref. 10). Hence the eigenvalue Q may be witten as

$$Q^2 = (j-1)(j+1) + 1/4 + 3/4 = j^2.\tag{85}$$

Since the value of the j is half a positive odd integer and $J^2 \geq 0$, we find

$$Q^2 = (3/2)^2, (5/2)^2, (7/2)^2, \cdots\cdots.\tag{86}$$

Thus we have the desired values of the Q as

$$Q = \pm 3/2, \pm 5/2, \pm 7/2, \cdots\cdots.\tag{87}$$

It should be noted that the eigenvalue of Q_0 is half an odd integer.

§ 5. A simple example of mass spectrum

We shall now study in this section the masses of particles as eigenvalues of an iterated Dirac equation by assuming a simple self-potential. All that would seem to be necessary is to consider the radial part. In order to find the radial part of the Dirac equation,

$$(\gamma_\mu \partial_\mu + \mu_0 + V(R))\psi = 0,\tag{88}$$

we have to discover the expression $(\gamma_\mu p_\mu)$ in terms of R and $\partial/\partial R$. This can be done by the method of Dirac.[12] We define a dynamical variable p_R as

$$p_R = 1/R \cdot (x_\mu p_\mu).\tag{89}$$

The true momentum conjugate to the variable R is

$$p_R - 3/2 \cdot i/R.\tag{90}$$

After some calculations we find

$$(\gamma_\mu x_\mu)(\gamma_\nu p_\nu) = (x_\mu p_\mu) + i(\sigma : L). \tag{91}$$

With the help of the relation (80) and (90) we get

$$(\gamma_\mu x_\mu)(\gamma_\nu p_\nu) = R\left(p_R - \frac{3}{2}\frac{i}{R}\right) + i\gamma_5 Q_0. \tag{92}$$

Introducing a matrix ϵ which satisfies the relations

$$R\epsilon = (\gamma_\mu x_\mu), \quad \epsilon^2 = 1, \quad \epsilon\gamma_5 + \gamma_5\epsilon = 0, \tag{93}$$

we obtain from (92)

$$(\gamma_\mu p_\mu) = \epsilon\left(p_R - \frac{3}{2}\frac{i}{R}\right) + \frac{i}{R}\epsilon\gamma_5 Q_0, \tag{94}$$

where use is made of the relations

$$\left.\begin{array}{ll}
[Q_0, (\gamma_\mu p_\mu)] = 0, & [Q_0, (\gamma_\mu x_\mu)] = 0, \\
[p_R, \epsilon] = 0, & [R, \epsilon] = 0, \\
[Q_0, \epsilon] = 0.
\end{array}\right\} \tag{95}$$

and

Thus the Dirac equation to be solved is

$$\left[\epsilon\left(\frac{d}{dR} + \frac{3}{2}\frac{1}{R}\right) - \frac{1}{R}\epsilon\gamma_5 Q + V(R) + \mu_0\right]\varphi(R) = 0. \tag{96}$$

Since the matrices ϵ and γ_5 satisfy the relation (93), it is convenient to adopt the representation of these matrices;

$$\epsilon = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \gamma_5 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{97}$$

Correspondingly, we put

$$\varphi(R) = \frac{1}{R^{3/2}}\begin{pmatrix} \phi_1(R) \\ \phi_2(R) \end{pmatrix}. \tag{98}$$

Then the equation (96) becomes

$$\left.\begin{array}{l}
d\phi_1/dR - Q/R \cdot \phi_1 + (V + \mu_0)\phi_2 = 0, \\
d\phi_2/dR + Q/R \cdot \phi_2 + (V + \mu_0)\phi_1 = 0.
\end{array}\right\} \tag{99}$$

The wave function is normalized,

$$\int_0^\infty (\phi_1^*\phi_1 - \phi_2^*\phi_2)\,dR = 1. \tag{100}$$

The introduction of new wave functions χ_1 and χ_2 as defined by

$$\left.\begin{array}{l}
\phi_1 = e^{-\mu_0 R}(\chi_1 - \chi_2), \\
\phi_2 = e^{-\mu_0 R}(\chi_1 + \chi_2),
\end{array}\right\} \tag{101}$$

then brings the wave equations (99) into the form

$$d\chi_1/dR + Q/R \cdot \chi_2 = -V\chi_1,$$
$$d\chi_2/dR + Q/R \cdot \chi_1 = (V + 2\mu_0)\chi_2. \tag{102}$$

Eliminating the component χ_2 from (102), we obtain a second-order differential equation

$$\frac{d^2\chi_1}{dR^2} + \left(\frac{1}{R} - 2\mu_0\right)\frac{d\chi_1}{dR} + \left(-2\mu_0 V - \frac{Q^2}{R^2} + \frac{dV}{dR} + \frac{V}{R} - V^2\right)\chi_1 = 0. \tag{103}$$

It is obvious that there is difficulty in carrying out the integration of the equation (103) because of the complicated form of the self-potential $V(R)$ which was shown in (46). Therefore some approximations will be made both for the equation (103) and for the self-potential $V(R)$. It is hoped that more general treatment will be given in a future paper. Here the emphasis will be on the derivation of a mass spectrum by means of reasonable approximations and the idealization of the model. Now, as to the self-potential itself, we assume that

$$V(R) \to 0, \qquad R \to \infty, \tag{104}$$

and, for simpliciy, $V(R)$ is of the Coulomb type. There are many ways of doing so, but it is better to examine the case of the expanded potential $V(R)$ under the assumption that the rest-mass m_0 is large compared to the mass correction. Therefore, replacing m in the expression (47), by m_0, we have the form

$$V(R) \sim -\frac{3g^2}{16\pi^2}\left(\frac{m}{m_0}\right)\exp\left(-\frac{m_0 R}{2}\right)\left(\frac{1}{R} - \frac{2}{m_0 R^2} + \frac{8}{m_0^2 R^3} - \cdots\right). \tag{105}$$

so that a Coulomb type self-potential is given by

$$V(R) = -g_0^2(m/m_0)1/R, \tag{106}$$

with

$$g_0^2 = 3/16\pi^2 \cdot g^2. \tag{107}$$

The reason for this choice is in the fact that (106) gives us an eigenvalue problem in which the masses are determined. Indeed, once another form for $V(R)$ is assumed, for example, $V(R)$ in which all m is replaced by m_0, the eigenvalue problem will lead to the quantization of the charge g.

By substituting the potential (106) into the equation (103), one might think that the Dirac equation could be solved exactly. We find, however, that there is no solution giving the physically acceptable masses of particles. Since the object of this method is to see what kind of mass spectra can be inferred under some assumptions, it becomes convenient at this stage to use a perturbation method to solve the iterated Dirac equation (103). Let us consider the terms which are proportional to $g^{2n}(n \geq 2)$ to be small perturbations. Then the perturbed wave function and mass level are given by

$$\chi_1 = \chi_1^{(0)} + \chi_1^{(1)} + \cdots, \tag{108}$$

and

$$m = m^{(0)} + m^{(1)} + \cdots .$$ (109)

We now find a solution for the $\chi_1^{(0)}$ in the form

$$\chi_1^{(0)}(R) = R^{\alpha}(a_0 + a_1 R + \cdots + a_\nu R^\nu + \cdots).$$ (110)

Since the constant a_0 cannot be zero, one finds

$$\alpha = |Q|.$$ (111)

From the condition that the series (110) must terminate, we get the mass levels as follows:

$$m^{(0)} = m_0/g_0^2 \cdot (n + |Q|),$$ (112)

where

$$n = 0, 1, 2, 3, \cdots\cdots,$$ (113)

and

$$|Q| = 3/2, 5/2, 7/2, \cdots\cdots .$$ (114)

Incidentally this result shows a mass spectrum which is similar to that given by Nambu[13] for the fermion. That is to say, according to him, the mass of the fermion except for that of the electron is expressed as half an odd integer in the units of Heisenberg in which $\hbar = c = r_0 = 1$ with the classical electron radius $r_0 = 2.80 \times 10^{-13}$ cm. However, we can say that at the present stage to compare the result with the experimental facts seems to be premature, because we have no rigorous proof of the mathematical consistency of the present treatment and have adopted the extremely simplified self-potential.

§ 6. Concluding remarks

The foregoing work provides a general theory of the mass spectrum. We have seen that, so far as we restrict ourselves to the mass-type divergences, there seems to be a possibility of avoiding the difficulty and at the same time of obtaining the mass spectra of particles, only if we assume an interaction which leads to negative self-energies and a rest-mass m_0 which may be a fundamental constant such as $1/r_0$. It is not a perfect theory, because it involves many assumptions. Nevertheless, there are interesting points presented by this idea. The first point is that, on the basis of this formulation, we can understand why the so-called mixed theories[4] are useful for the self-energy problem. As is easily seen, a positive self-energy, for instance, the electromagnetic self-energy of the Dirac particle does not give us discrete mass levels just as in the case of the energy levels for the Coulomb repulsive potential in the elementary quantum mechanics. Hence, it may be necessary to mix several fields in order to get mass spectrum for the following two reasons. One is that we may need not to cancel the divergences of the self-energy but to make the total self-energy negative by considering various fields. Taking into account this point, one may say that the mass relation of heavy particles, which was derived by means of the compensation method,[4] shows the limiting mass values of particles involved.

The other is that, in the case that there remain such higher singularities as quadratic divergences in the self-potential, it may be required to eliminate them in order that one can solve the wave equation for the system. This is the same procedure as that of Møller and Rosenfeld[14] in the theory of nuclear forces. At any rate, the above discussion reveals that it seems to be necessary to take the cohesive fields into consideration for the analysis of the mass problem. The second point is that there has long been feeling that the states of unstable heavy particles may correspond to the higher angular momentum states of the elementary particles. Such an idea was developed so far. Our method shows in this context how the eigenvalue of the generalized angular momentum can be related to the mass levels.

Finally, we may imagine a model of the elementary particle in the following way; the particle is embedded in its fields in such a manner as to produce negative self-potentials which give us discrete mass levels. The decay of particles may indicate a transition from a state to the other state, which is caused by some other interactions. Thus it must be stressed that the space-time point of view may furnish a new interpretation of the problems of the mass spectra and divergence difficulties.

In conclusion, I should like to thank Prof. H. Yukawa and Prof. G. Yevick for their kind discussions, the Fulbright Foundation in Japan for a grant, and the Department of Physics at Columbia University for its hospitality.

References

1) Proceedings of the Third Annual Rochester Conference, 1952.
2a) H. Yukawa, Phys. Rev. **77** (1950), 218 ; **80** (1950), 1047 ; **91** (1953), 415, 416.
2b) A. Pais, Physica **19**(1953), 869 ; Theor. Phys. **10** (1953), 457.
3) J. Schwinger, Phys. Rev. **74** (1948), 1439 ; **75** (1949), 651 ; **76** (1949), 790.
4) H. Enatsu, Prog. Theor. Phys. **6** (1951), 643 ; Phys. Rev. **85** (1952), 483 ; **89** (1953), 1304.
5) P. Kristensen and C. Møller, Kgl. Dansk. Vidensk. Proc. **27** (1952), No. 7.
     C. Bloch, ibid. **27** (1952), No. 8.
     A. Pais and G. E. Uhlenbeck, Phys. Rev. **79** (1950), 145.
     W. Pauli, Nuovo Cimento, **10** (1953), 648.
     R. E. Peierls and H. MacManus, Proc. Roy. Soc. A **195** (1948), 323.
     O. Hara and T. Marumori, Prog. Theor. Phys. **9** (1953), 559.
     Y. Katayama, Prog. Theor. Phys. **10** (1953), 31.
6) J. Schwinger, Phys. Rev. **82** (1951), 664.
     R. P. Feynman, **80** (1950), 440 ; **84** (1951), 108.
     G. Falk, Zeits. f. Phys. **132** (1952), 44.
7) P.A.M. Dirac, *Principles of quantum mechanics* (1930), First edition.
8) M. Neuman, Phys. Rev. **85** (1952), 129.
9) S. Tomonaga, Prog. Theor. Phys. **1** (1946), 27.
10) M. Born and K. Fuchs, Proc. Roy. Soc. Edin. **60** (1940), 100, 141.
     K. Fuchs, Proc. Roy. Soc. Edin. **60** (1940), 147.
11) L. Infeld and T.E. Hull, Rev. Mod. Phys. **23** (1951), 21.
12) P.A.M. Dirac, *Principles of quantum mechanics* (1947), Third edition, p. 266.
13) Y. Nambu, Prog. Theor. Phys. **7** (1952), 595.
14) C. Møller and L. Rosenfeld, Kgl. Dansk. Vidensk. Proc. **17** (1940), No. 8.

Generalized Racah Coefficient and its Applications

Akito ARIMA, Hisashi HORIE and Yukito TANABE

Department of Physics, University of Tokyo

Generalized Racah coefficient, designated as U coefficient in this paper, has been defined as the transformation function between two different coupling schemes in pairs of any four angular momenta, corresponding to the Racah coefficient defined as the transformation function between two different coupling schemes of any three angular momenta. Several simple properties of the U coefficient have been derived, and the method of tensor operators made to be extended to more general problems. Transformation coefficients between LS- and jj-coupling schemes in a many particle system can be evaluated by making use of these coefficients.

§ 1.   Introduction

The Racah coefficient has proved to play a very important role in detailed theories of the atomic and nuclear spectroscopy[1-3], and also to be useful for the studies of the nuclear radiations and reactions.[4),5)] It is defined as the transformation function between two different coupling schemes of any three angular momenta j_1, j_2 and j_3 by

$$[(2J_{12}+1)(2J_{23}+1)]^{1/2}W(j_1j_2Jj_3;\ J_{12}J_{23}) = (j_1j_2(J_{12})j_3J\,|\,j_1,j_2j_3(J_{23})J), \quad (1)$$

where $j_1+j_2=J_{12},\ j_2+j_3=J_{23}$ and $J_{12}+j_3=J_1+J_{23}=J$. In a similar way, we can define the generalized Racah coefficient which is designated as the U coefficient in this paper, as the transformation function between two different coupling orders in pairs of any four angular momenta j_1,j_2,j_3 and j_4 by

$$[(2J_{12}+1)(2J_{34}+1)(2J_{13}+1)(2J_{24}+1)]^{1/2}U\begin{pmatrix} j_1 & j_2 & J_{12} \\ j_3 & j_4 & J_{34} \\ J_{13} & J_{24} & J \end{pmatrix}$$

$$= (j_1j_2(J_{12})j_3j_4(J_{34})J\,|\,j_1j_3(J_{13})j_2j_4(J_{24})J), \quad (2)$$

where $j_1+j_2=J_{12}, j_3+j_4=J_{34}, j_2+j_4=J_{24}, j_1+j_3=J_{13}$ and $J_{12}+J_{34}=J_{13}+J_{24}=J$, and the nine angular momenta as the arguments in the U coefficients are arranged in three rows and columns in natural order. It is, therefore, expressed by a sum of the products of six Clebsch-Gordan coefficients of the vector additions as

$$(j_1j_2(J_{12})j_3j_4(J_{34})J\,|\,j_1j_3(J_{13})j_2j_4(J_{24})J) = \sum(j_1j_2\,m_1m_2\,|\,j_1j_2J_{12}M)$$

$$\cdot\,(j_3j_4\,m_3m_4\,|\,j_3j_4J_{34}\,M_{34})(J_{12}J_{34}\,M_{12}\,M_{34}\,|\,J_{12}J_{34}\,JM)(j_1j_3\,m_1m_3\,|\,j_1j_3\,J_{13}M_{13})$$

$$\cdot\,(j_2j_4\,m_2\,m_4\,|\,j_2j_4\,J_{24}\,M_{24})(J_{13}J_{24}\,M_{13}\,M_{24}\,|\,J_{13}\,J_{24}JM), \quad (3)$$

where the summation is extended over all possible values of m_1, m_2, m_3, m_4, M_{12}, M_{34}, M_{13} and M_{24}, restricted by obvious relations $m_1 + m_2 = M_{12}$, $m_3 + m_4 = M_{34}$, $m_1 + m_3 = M_{13}$, $m_2 + m_4 = M_{24}$, and $M_{12} + M_{34} = M_{13} + M_{24} = M$.

It is shown in the next section, that the U coefficient can be expressed in terms of a sum of the products of three Racah coefficients and that the Racah coefficient is a special case of the U coefficient. Some other properties of the coefficient, also, will be derived there. In sec. 3, the method of tensor operators are extended to more general operators which are constructed as tensor products of two tensor operators. It enables us to treat the spin-dependent interactions in a more general way. And in sec. 4, the method of the calculation of the transformation function between LS- and jj-coupling schemes is derived which seems very important in the treatment of nuclear shell model, especially for light nuclei. Finally some recurrence formulae for the coefficients are given in the Appendix.

§ 2. Properties of the U coefficients

The U coefficient in (2) is defined for integral and half-integral values of the nine parameters, with the limitation that each of the six triads

$$(j_1, j_2, J_{12}), \quad (j_3, j_4, J_{34}), \quad (J_{12}, J_{34}, J), \quad (j_1, j_3, J_{13}), \quad (j_2, j_4, J_{24}), \quad (J_{13}, J_{24}, J) \quad (4)$$

has an integral sum, and vanishes unless the elements of each triad (4) satisfy the triangular inequalities according to the definition (2) and (3).

The summation in (3) can be carried out by making use of the Racah coefficients if we introduce the following intermediate state characterised by $J_{123} = J_{12} + j_3$, that is

$$(j_1 j_2 (J_{12}) j_3 j_4 (J_{34}) J \,|\, j_1 j_3 (J_{13}) j_2 j_4 (J_{24}) J)$$
$$= \sum_{J_{123}} (J_{12}, j_3 j_4 (J_{34}) J \,|\, J_{12} j_3 (J_{123}) j_4 J) (j_1 j_2 (J_{12}) j_3 J_{123} \,|\, j_1 j_3 (J_{13}) j_2 J_{123})$$
$$\cdot (J_{13} j_2 (J_{123}) j_4 J \,|\, J_{13}, j_3 j_4 (J_{34}) J), \quad (5)$$

where the summation over J_{123} is extended over all possible values compatible with the condition $J_{123} = J_{12} + j_3 = J_{13} + j_2$. Therefore the U coefficient can be expressed in terms of Racah coefficients with RIII (4) and (5). In abbreviated notations for arguments, it is given by

$$U \begin{pmatrix} a & b & c \\ c & d & e' \\ f & f' & g \end{pmatrix} = \sum_{\lambda} (2\lambda + 1) W(fgbd; f'\lambda) W(cgcd; e'\lambda) W(fcbe; a\lambda). \quad (6)$$

It is easily seen that the Racah coefficient can be obtained as a special case of the U coefficient in which any one of the six arguments b, c, d, e, f and g appearing in any two W's in the right hand side is equal to zero. For example, if $g = 0$, $c = c'$ and $f = f'$ result for non-vanishing U, which is given by

$$U\begin{pmatrix} a & b & c \\ c & d & e \\ f & f & 0 \end{pmatrix}=(-1)^{e+f-a-d}\,W(abcd;\,ef)/[\,(2e+1)\,(2f+1)\,]^{1/2}. \tag{7}$$

Owing to the symmetry properties of the U coefficient which we can show in the following, the coefficient reduces always to the W coefficient if any one of its nine arguments is equal to zero.

We can immediately derive the following symmetry properties from those of the Racah coefficients (see RII (40a) and (40b)) and the relations given by RII (43) and Biedenharn, Blatt and Rose's[6] (17).

i) Transposition of " rows " and " columns " :

$$U\begin{pmatrix} a & b & e \\ c & d & e' \\ f & f' & g \end{pmatrix}=U\begin{pmatrix} a & c & f \\ b & d & f' \\ e & e' & g \end{pmatrix}. \tag{8}$$

ii) Interchanges of two " rows " or " columns " ;

$$U\begin{pmatrix} a & b & e \\ c & d & e' \\ f & f' & g \end{pmatrix}=(-1)^{\sigma}U\begin{pmatrix} c & d & e' \\ a & b & e \\ f & f' & g \end{pmatrix}=(-1)^{\sigma}U\begin{pmatrix} f & f' & g \\ c & d & e' \\ a & b & e \end{pmatrix}, \tag{9}$$

where $\sigma=a+b+c+d+e+e'+f+f'+g$ $(=\text{integer})$.

Combining (8) and (9), we obtain 72 different arrangements of the nine parameters. For example, we can rewrite formula (6) into the following more symmetrical form :

$$U\begin{pmatrix} a & b & e \\ c & d & e' \\ f & f' & g \end{pmatrix}=(-1)^{\sigma}\sum_{\lambda}(2\lambda+1)\,W(bcf;\,\lambda a)\,W(bcf'e';\,\lambda d)\,W(efe'f';\,\lambda g), \tag{10}$$

where the " diagonal " elements of U appear as the last arguments of the three W coefficients. We shall prefer this form as the standard formula connecting the U and the W coefficients. Furthermore, it is easy to see that, if $a=c$, $b=d$ and $e=e'$, according to (9),

$$U\begin{pmatrix} a & b & e \\ a & b & e \\ f & f' & g \end{pmatrix}=0, \qquad (f+f'+g=\text{odd}). \tag{11}$$

Some identities can be derived from the definition of the U coefficient as a transformation function between two different couplings in pairs of four angular momenta. Four angular momenta a, b, c and d can be combined into various pairs, as follows :

$$g=(a+b)+(c+d)=(a+c)+(b+d)=(a+d)+(b+c). \tag{12}$$

It follows at once that

$$\sum_{ff'}(ab(e)cd(e')g \mid ac(f)bd(f')g)(ac(f)bd(f')g \mid ab(e_1)cd(e_1')g)$$
$$= \delta(e, e_1)\delta(e', e_1'). \tag{13}$$

In terms of U's, this gives us from (2), the following orthogonality relation between them:

$$\sum_{ff'}(2f+1)(2f'+1)U\begin{pmatrix} a & b & e \\ c & d & e' \\ f & f' & g \end{pmatrix}U\begin{pmatrix} a & b & e_1 \\ c & d & e_1' \\ f & f' & g \end{pmatrix}$$
$$= \delta(e, e_1)\delta(e', e_1')/[(2e+1)(2e'+1)]. \tag{14}$$

Since the transformation function between the first and third coupling orders in pairs of (12) can be expressed, through the second one, as

$$\sum_{ff'}(ab(e)cd(e')g \mid ac(f)bd(f')g)(ac(f)bd(f')g \mid ad(h)bc(h')g)$$
$$= (ab(e)cd(e')g \mid ad(h)bc(h')g), \tag{15}$$

we obtain, another useful relation between U's :*

$$\sum_{ff'}(-1)^{e'-f'-2c+h'}(2f+1)(2f'+1)U\begin{pmatrix} a & b & e \\ c & d & e' \\ f & f' & g \end{pmatrix}U\begin{pmatrix} a & c & f \\ d & b & f' \\ h & h' & g \end{pmatrix}=U\begin{pmatrix} a & b & e \\ d & c & e' \\ h & h' & g \end{pmatrix}. \tag{16}$$

We can see without difficulties that (15) and (16) are the generalization of RII (42) and (43), if we put g equal to zero. Beside these relations we have obtained several other relations between U's and between U's and H''s, some of which will be given in Appendix. We shall also show there how recurrence formulae for U coefficients are obtained from one of them.

§ 3.   Application to the calculation of matrix elements
of tensor operators

(a)   Tensor product of two tensor operators

The tensor product of two tensor operators $T^{(k_1)}$ and $U^{(k_2)}$ is defined in the usual way by an irreducible form

---

*)   The following relations are easily proved:

$$(ab(e)cd(e')g \mid ad(f)bc(f')g) = (-1)^{c+d-e'}[(2e+1)(2e'+1)(2f+1)(2f'+1)]^{1/2}U\begin{pmatrix} a & b & e \\ d & c & e' \\ f & f' & g \end{pmatrix},$$

and

$$(ab(e)cd(e')g \mid ad(f)cb(f')g) = (-1)^{d-e'-b+f'}[(2e+1)(2e'+1)(2f+1)(2f'+1)]^{1/2}U\begin{pmatrix} a & b & e \\ d & c & e' \\ f & f' & g \end{pmatrix}.$$

$$[\boldsymbol{T}^{(k_1)} \times \boldsymbol{U}^{(k_2)}]_Q^{(K)} = \sum_{q_1 q_2} T_{q_1}^{(k_1)} \cdot U_{q_2}^{(k_2)} (k_1 k_2 q_1 q_2 \mid k_1 k_2 KQ). \tag{17}$$

In practical applications the most important tensor products are those in which two tensor operators operate on different parts of a composite system. The operator of this type appears in many problems, for example, in the calculation of matrices of spin-dependent interactions[7-9], of multipole moments of radiations in the nuclear shell model[10], and of polarization of emerging particles in nuclear reactions[11].*

When $\boldsymbol{T}^{(k_1)}$ operates on system 1 and $\boldsymbol{U}^{(k_2)}$ operates on system 2, the matrix element of a tensor product of $\boldsymbol{T}^{(k_1)}$ and $\boldsymbol{U}^{(k_2)}$ in $(j_1 j_2 JM)$ scheme is given by

$$(j_1 j_2 JM \mid [\boldsymbol{T}^{(k_1)} \times \boldsymbol{U}^{(k_2)}]_Q^{(K)} \mid j_1' j_2' J'M') = \sum (j_1 j_2 JM \mid j_1 j_2 m_1 m_2)$$

$$\cdot (j_1 m_1 \mid T_{q_1}^{(k_1)} \mid j_1' m_1')(j_2 m_2 \mid U_{q_2}^{(k_2)} \mid j_2' m_2')$$

$$\cdot (j_1' j_2' m_1' m_2' \mid j_1' j_2' J'M')(k_1 k_2 q_1 q_2 \mid k_1 k_2 KQ). \tag{18}$$

And, if the double-barred element is defined in accordance with RII (29) by

$$(\alpha j m \mid T_q^{(k)} \mid \alpha' j'm') = (\alpha j \| T^{(k)} \| \alpha' j)(j'km'q \mid j'kjm)/(2j+1)^{1/2}, \tag{19}$$

it is easy to be shown that the double-barred elements of the tensor product $[\boldsymbol{T}^{(k_1)} \times \boldsymbol{U}^{(k_2)}]^{(K)}$ are expressed in terms of those of $\boldsymbol{T}^{(k_1)}$ and $\boldsymbol{U}^{(k_2)}$ and a U coefficient as follows :

$$(j_1 j_2 J \| [T^{(k_1)} \times U^{(k_2)}]^{(K)} \| j_1' j_2' J) = (j_1 \| T^{(k_1)} \| j_1)(j_2 \| U^{(k_2)} \| j_2')$$

$$\cdot [(2J+1)(2J'+1)(2K+1)]^{1/2} U \begin{pmatrix} j_1 & j_1' & k_1 \\ j_2 & j_2' & k_2 \\ J & J' & K \end{pmatrix}. \tag{20}$$

This simple and symmetrical formula is a natural generalization of RII (38), (44a) and (44b). First of all, noting that

$$(\boldsymbol{T}^{(k)} \cdot \boldsymbol{U}^{(k)}) = (-1)^k (2k+1)^{1/2} [T^{(k)} \times U^{(k)}]_0^{(0)}, \tag{21}$$

where in the left-hand side $(\boldsymbol{T}^{(k)} \cdot \boldsymbol{U}^{(k)}) = \sum_q (-1)^q T_q^{(k)} U_{-q}^{(k)}$ represents the scalar product of the two tensor operators $\boldsymbol{T}^{(k)}$ and $\boldsymbol{U}^{(k)}$, and putting $k_1 = k_2 = k$ and $K = 0$ in (20), we obtain the relation RII (38)

$$(j_1 j_2 JM \mid (\boldsymbol{T}^{(k)} \cdot \boldsymbol{U}^{(k)}) \mid j_1' j_2' JM) = (-1)^{j_1 + j_2' - J}$$

$$\cdot (j_1 \| T^{(k)} \| j_1')(j_2 \| U^{(k)} \| j_2') W(j_1 j_2 j_1' j_2' ; Jk). \tag{22}$$

Putting further $k_2 = 0$ and $k_1 = K = k$ in (20), and noting that $(j \| 1 \| j) = (2j+1)^{1/2}$, we get the relation given by RII (44a)

$$(j_1 j_2 J \| T^{(k)} \| j_1' j_2' J') = (-1)^{k+j_2-j_1'-J}$$
$$\cdot (j_1 \| T^{(k)} \| j_1')[(2J+1)(2J'+1)]^{1/2} W(j_1 J j_1' J'; j_2 k). \qquad (23)$$

The relation RII (44b) can be obtained in a similar way by putting $k_1=0$ and $k_2=K=k$. It must be observed that for the double-barred elements of the tensor product,

$$(j_1 j_2 J \| [T^{(k_1)} \times U^{(k_2)}]^{(K)} \| j_1' j_2' J') = (-1)^{J-J''-k_1-k_2-K}$$
$$\cdot (j_1' j_2' J' \| [T^{(k_1)} \times U^{(k_2)}]^{(K)} \| j_1 j_2 J), \qquad (24)$$

corresponding to RII (31). As a special case of (24), the following formula is obtained, by putting $j_1=j_1'$, $j_2=j_2'$ and $J=J'$,

$$(j_1 j_2 J \| [T^{(k_1)} \times U^{(k_2)}]^{(K)} \| j_1 j_2 J) = 0 \qquad (k_1+k_2-K=\text{odd}), \qquad (25)$$

which can also be derived from (11) immediately.

(b) Matrix elements of the scalar product of two irreducible composite tensors

In order to calculate the matrix of spin-dependent interactions, it is necesary to treat the scalar product of two irreducible composite tensors. Hence we consider this quantity more in detail. Let $T^{(k_1 k_2: K)}$ be an irreducible composite tensor of degree K with respect to $J=j_1+j_2$ which behaves as an irreducible tensor of degree k_1 and k_2 with respect to j_1 and j_2 respectively (This may be considered as an abbreviation of the tensor product given in (17)), and $U^{(k_1 k_2: K)}$ have a similar meaning. By making use of (20) and BBR (1), the matrix elements of this scalar product in $(j_1 j_2 JM)$ scheme are given by

$$(\gamma j_1 j_2 JM | (T^{(k_1 k_2: K)} \cdot U^{(k_1 k_2: K)}) | \gamma j_1' j_2' JM)$$
$$= (2K+1) \sum_{\gamma'' j_1'' j_2''} (\gamma j_1 j_2 \| T^{(k_1 \ k_2)} \| \gamma'' j_1'' j_2'')(\gamma'' j_1'' j_2'' \| U^{(k_1 \ k_2)} \| \gamma' j_1' j_2')$$
$$\cdot \sum_{J''} (-1)^{J-J''}(2J''+1) \ U \begin{pmatrix} j_1 & j_1'' & k_1 \\ j_2 & j_2'' & k_2 \\ J & J'' & K \end{pmatrix} U \begin{pmatrix} j_1'' & j_1' & k_1' \\ j_2'' & j_2' & k_2' \\ J'' & J & K \end{pmatrix}. \qquad (26)$$

The summation over J'' in (26) can be carried out, using the relation between U and W coefficients and with RII (43) and BBR (17), so that (26) is written in an expected form as

$$(\gamma j_1 j_2 JM | (T^{(k_1 k_2: K)} \cdot U^{(k_1 k_2: K)}) | \gamma' j_1' j_2' JM)$$
$$= (2K+1) \sum_{\lambda} (-1)^{j_1+j_2'-J} W(j_1 j_2 j_1' j_2'; J\lambda)$$
$$\cdot \sum_{\gamma'' j_1'' j_2''} (-1)^{k_1'+k_2} (\gamma j_1 j_2 \| T^{(k_1 \ k_2)} \| \gamma'' j_1'' j_2'')(\gamma'' j_1'' j_2'' \| U^{(k_1', k_2')} \| \gamma' j_1' j_2')$$
$$\cdot W(j_1 j_2' k_1 k_2'; \lambda j_1'') W(j_2 j_2' k_2 k_3'; \lambda j_2'') W(k_1 k_2 k_1' k_2'; K\lambda). \qquad (27)$$

This formula is useful, especially, in the treatment of the spin-dependent interactions[7-9],

in which this reduction has been done in a more straightforward way. For example, in the case of the spin-spin interaction between electrons, the angular momenta j_1 and j_2 are the total spin and the total orbital angular momenta respectively, and $k_1 = k_1' = 1$ and $k_2' = k_2 + 2$ $(k = 0, 2, \cdots)$, so that only $\lambda = 2$ appears in this equation.

(c) Coefficients of the exchange integrals of a many particle system

Racah has given a general method for obtaining the coefficient of exchange integrals in the case of electrostatic interactions RII, sec. 5. We shall extend this method to spin-dependent interactions, making use of the result in this section. The spin-dependent interaction can be represented as a scalar product of two irreducible tensors which have degree K ($K \neq 0$) in the spin and the ordinary space respectively; for example, in the tensor or the spin-spin interaction $K = 2$ and in the spin-orbit interaction $K = 1$.

First, we consider only the orbital part and assume that the irreducible tensor is a tensor product of two tensor operators of degree k_1 and k_2, the former operating on particle 1 and the latter on particle 2. Then we obtain the orbital part of coefficients of exchange integrals in terms of double-barred elements as

$$(-1)^{l_1 + l_2 - L'} (l_1 l_2 L \| [T_1^{(k_1)} \times U_2^{(k_2)}]^{(K)} \| l_2 l_1 L')$$
$$= (-1)^{l_1 + l_2 - L'} (l_1 \| T^{(k_1)} \| l_2)(l_2 \| U^{(k_2)} \| l_1)[(2L+1)(2L'+1)(2K+1)]^{1/2}$$
$$\cdot U \begin{pmatrix} l_1 & l_2 & L \\ l_1 & l_2 & L' \\ k_1 & k_2 & K \end{pmatrix},$$

and owing to (15) and RII (31), it follows that

$$(-1)^{l_1 + l_2 - L} (l_1 l_2 L \| [T^{(k_1)} \times U^{(k_2)}]^{(K)} \| l_2 l_1 L')$$
$$= (l_1 \| T^{(k_1)} \| l_2)(l_1 \| U^{(k_2)} \| l_2) [(2L+1)(2L'+1)(2K+1)]^{1/2}$$
$$\cdot \sum_{r,s}{}'(-1)^{k_2 - s} \cdot (2r+1)(2S+1) U \begin{pmatrix} l_1 & l_1 & r \\ l_2 & l_2 & s \\ k_1 & k_2 & K \end{pmatrix} U \begin{pmatrix} l_1 & l_2 & L \\ l_1 & l_2 & L' \\ r & s & K \end{pmatrix}. \quad (28)$$

Futhermore, if we define the unit tensor $u^{(k)}$ by

$$(l \| u^{(k)} \| l') = \delta(l, l'), \quad (29)$$

and take (20) into account, we may also write

$$(-1)^{l_1 + l_2 - L} (l_1 l_2 L \| [T^{(k_1)} \times U^{(k_2)}]^{(K)} \| l_2 l_1 L')$$
$$= (l_1 \| T^{(k_1)} \| l_2)(l_1 \| U^{(k_2)} \| l_2) \sum_{r,s}(-1)^{k_2 - s}(2r+1)(2s+1)$$
$$\cdot U \begin{pmatrix} l_1 & l_1 & r \\ l_2 & l_2 & s \\ k_1 & k_2 & K \end{pmatrix} \cdot (l_1 l_2 L \| [u_1^{(r)} \times u_2^{(s)}]^{(K)} \| l_1 l_2 L'). \quad (30)$$

Therefore, the calculation of the coefficients of the exchange integrals can be carried out in the same way as for obtaining those of direct integrals with respect to the operator

$$(l_1 \| T^{(k_1)} \| l_2)(l_1 \| U^{(k_2)} \| l_2) \sum_{r,s} (-1)^{k_2-s}(2r+1)(2s+1)$$

$$\cdot U \begin{pmatrix} l_1 & l_1 & r \\ l_2 & l_2 & s \\ k_1 & k_2 & K \end{pmatrix} [u_1^{(r)} \times u_2^{(s)}]^{(K)}, \tag{31}$$

in place of $|T^{k_1} \times U^{k_2}|$. In a similar way, another irreducible tensor of degree K can be obtained as the spin part, which becomes usually much simpler. Therefore, the complete operator necessary for the calculation of coefficients of exchange integrals is given by contructing scalar products of these two irreducible tensors of degree K with respect to the ordinary and spin spaces and reversing the total sign due to the antisymmetry of the wave functions.

As a trivial example of this procedure, the formula RII (59) of the coefficient $\eta_k(l_1 l_2 L)$ for the electrostatic interaction is derived by putting $K=0$ and $T^{k_1} = U^{k_2} = C^{(k)}$ in (31) where $C_q^{(k)} = [4\pi/(2k+1)]^{1/2} \theta(kq) \Phi(q)$. Another trivial example is given by the construction of Dirac's exchange operator with (30). Letting $l_1 = l_2 = 1/2$, $T^0 = U^{(0)} = 1$, and noting that $\mathbf{1} = (2)^{1/2} u^{(0)}$ and $\mathbf{s} = (3/2)^{1/2} u^{(1)}$, we immediately have Dirac's exchange operator

$$(-1)^{1-s} = 1/2 \cdot [1 + 2(\mathbf{s}_2 \cdot \mathbf{s}_1)]. \tag{32}$$

The simplest example of the spin-dependent interactions is given by the spin-spin interaction between electrons.[7] In this case, we need not change the form of the spin part since it is symmetrical with respect to the spin variables of two electrons. This holds also for the tensor interaction with arbitrary radial dependence. The coefficients of exchange integrals of the spin-spin interaction between electrons

$$\int \int R_{l_1}(r_1) R_{l_2}(r_2) \frac{r_1^k}{r_1^{k+3}} R_{l_1}(r_2) R_{l_2}(r_1) dr_1 dr_2$$

are given by the matrix elements of the operator

$$-2f_k(l_1 \| C^{(k)} \| l_2)(l_1 \| C^{k+2} \| l_2) \sum_{r,s}' (-1)^{k-s}(2r+1)(2s+1)$$

$$\cdot U \begin{pmatrix} l_1 & l_1 & r \\ l_2 & l_2 & s \\ k & k+2 & 2 \end{pmatrix} ([\mathbf{s}_1 \times \mathbf{s}_2]^{(2)} \cdot [u_1^{(n)} \times u_2^{(s)}]^{(2)}), \tag{33}$$

where the prime on the summation symbol denotes that the summation is extended only over those values of r and s which satisfy $r+s=$ even and the coefficients f_k is given by $(-1)^{k+1} 4 [(k+1)(k+2)(2k+1)(2k+3)(2k+5)/5]^{1/2}$.

§ 4. Transformation coefficients between *LS*- and *jj*-coupling schemes in l^n configuration

The transformation coefficient between *LS*- and *jj*-coupling schmes of two particle system can be obtained at once from the general expression for the transformation function between two different couplings in pairs of four angular momenta given in sec. 1 and 2. For two equivalent particles (identical particles which are contained in the same shell), however, the formula does not hold without modification on account of the Pauli exclusion principle as will be seen in the following. Here we consider the case in which there are n equivalent particles in the same shell with azimuthal quantum number l. The states of l^n configuration are characterized by $\alpha SLJM$ in *LS*-oupling scheme, where α is the quantum number other than S, L, J and M. On the other hand, the states with the same J and M are characterized by $j_1^{n_1}(\beta_1 J_1)j_2^{n_2}(\beta_2 J_2)JM$ in *jj*-coupling scheme, where $j_1 = l+1/2$, $j_2 = l-1/2$ and $n_1 + n_2 = n$, β's being the quantum number other than J and M. The quantum number of the isotopic spin employed sometimes in nuclear shell model can be included in α and in β. Taking into account the antisymmetry property of wave functions, the transformation function between *LS*- and *jj*-coupling schemes for n equivalent particles can be obtained in terms of those for $(n-1)$ equivalent particles, the coefficients of fractional parentages, and the U and W coefficients :

$$(l^n \alpha S L\, JM \,|\, j_1^{n_1}(\beta_1 J_1)j_2^{n_2}(\beta_2 J_2)\,JM)$$

$$= (-1)^{n_2}(n_1/n)^{1/2}\sum (l^n \alpha SL\{| l^{n-1}(\alpha'S'L')lsL)\left(S'\tfrac{1}{2}(S)L'l(L)J\,|\,S'L'(J')\tfrac{1}{2}l(j_1)J\right)$$

$$\cdot (l^{n-1}\alpha'S'L'J'\,|\,j_1^{n_1-1}(\beta_1'J_1')j_2^{n_2}(\beta_2 J_2)J')(J_1'J_2(J')j_1 J\,|\,J_1'j_1(J_1)J_2 J)$$

$$\cdot(j_1^{n_1-1}(\beta_1'J_1')\,j_1 J_1|\} j_1^{n_1}\beta_1 J_1) + (n_2/n)^{1/2}\sum(l^n\alpha SL\{| l^{n-1}(\alpha'S'L')lSL)$$

$$\cdot\left(S'\tfrac{1}{2}(S)L'l(L)J\,|S'L'(J')\tfrac{1}{2}l(j_2)J\right)(l^{n-1}\alpha'S'L'J'\,|\,j_1^{n_1}(\beta_1 J_1)j_2^{n_2-1}(\beta_2'J_2')J')$$

$$\cdot(J_1 J_2'(J')j_2 J\,|\,J_1,\,J_2'j_2(J_2)J)(j_2^{n_2-1}(\beta_2'J_2')\,j_2 J_2|\}j_2^{n_2}\beta_2 J_2). \tag{34}$$

The tables of the coefficients of fractional parentages were given by RIII for the atomic p^n and d^n configurations, by Jahn and van Wieringen[2] for the nuclear p^n, by Jahn[2] for d^3 and d^4 configurations in *LS*-coupling, and Edmond and Flowers[3] for $(3/2)^3$, $(3/2)^4$, $(5/2)^5$ $(7/2)^3$ and$(7/2)^4$ configurations. The tables of W coefficients were given by Biedenharn[12] and Obi *et al.*[13]

For two equivalent particles which have only states with $S+L=$ even in *LS*-coupling scheme, eq. (34) reduces to

$$(l^2 SLJM\,|\,j^2\,JM) =\left(\tfrac{1}{2}\,\tfrac{1}{2}\,(S)ll(L)J\,\Big|\,\tfrac{1}{2}l(j)\tfrac{1}{2}\,l(j)J\right),\qquad j=j_1 \text{ or } j_2,$$

$$(l^2 SLJM\,|\,j_1 j_2 JM) = (2)^{1/2}\left(\tfrac{1}{2}\,\tfrac{1}{2}\,(S)ll(L)J\,\Big|\,\tfrac{1}{2}\,l(j_1)\tfrac{1}{2}\,l(j_2)J\right).$$

Appendix. Identities and recurrence formulae for the U coefficients

The values of the U coefficients can be obtained by inserting the values of the W coefficients in the formula (6). However, there are also some recurrence formulae between the U coefficients which may be available for the evaluation of the coefficients.

In order to obtain an identity from which recurrence formulae can be derived, we consider the following two different coupling schemes of five angular momenta and the transformation function between them. It is evident that

$$(j_1 j_2 (J_{12}) j_3 j_4 (J_{34})(J_{1234}) j_5 J \mid j_1 j_3 (J_{13}), j_2 j_4 (J_{24}) j_5 (J_{245}) J)$$
$$= (j_1 j_2 (J_{12}) j_3 j_4 (J_{34}) J_{1234} \mid j_1 j_3 (J_{13}) j_2 j_4 (J_{24}) J_{1234})$$
$$\cdot (J_{13} J_{24} (J_{1234}) j_5 J \mid J_{13}, J_{24} j_5 (J_{245}) J). \tag{A. 1}$$

But this transformation function can be expressed in another way by employing two intermediate state as

$$\sum_{J_{345} J_{45}} (J_{12} J_{34} (J_{1234}) j_5 J \mid J_{12}, J_{34} j_5 (J_{345}) J)(j_3 j_4 (J_{34}) j_5 J_{345} \mid j_3, j_4 j_5 (J_{45}) J_{345})$$
$$\cdot (j_1 j_2 (J_{12}) j_3 J_{45} (J_{345}) J \mid j_1 j_2 (J_{12}) j_2 J_{45} (J_{245}) J)$$
$$\cdot (j_2 j_4 j_5 (J_{45}) J_{245} \mid j_2 j_4 (J_{24}) j_5 J_{245}). \tag{A. 2}$$

Equating this with (A. 1) and expressing the result by the U and W coefficients, we obtain the relation between them as follows ;

$$U\begin{pmatrix} a & b & e \\ c & d & e' \\ f & f' & g \end{pmatrix} W(ff'\bar{g}h ; g\bar{f}') = \sum_{\lambda\mu}(2\lambda+1)(2\mu+1) W(ee'\bar{g}h ; g\lambda)$$

$$\cdot W(cd\lambda h ; e'\mu) W(bd\bar{f}'h ; f'\mu) U\begin{pmatrix} a & b & e \\ c & \mu & \lambda \\ f & f' & g \end{pmatrix}. \tag{A. 3}$$

Applications of this identity to give recurrence formulae are immediate. Take, for example, $h = 1\ 2$. Then in the summation on the right hand side of (A. 3) λ and μ take only two values $\lambda = e' \pm 1/2$ and $\mu = d \pm 1/2$ respectively. Then, we can choose the values of $\bar{f}'$ and $\bar{g}$ as $\bar{f}' = f' \pm 1/2$ and $\bar{g} = g \pm 1/2$. Therefore, the values of

$$U\begin{pmatrix} a & b & e \\ c & d-1/2 & e'-1/2 \\ f & f'-1/2 & g-1/2 \end{pmatrix},$$ for example, can be evaluated in terms of four U coefficients

$$U\begin{pmatrix} a & b & e \\ c & d & e' \\ f & f' & g \end{pmatrix}, \quad U\begin{pmatrix} a & b & e \\ c & d & e'-1 \\ f & f' & g \end{pmatrix}, \quad U\begin{pmatrix} a & b & e \\ c & d-1 & e' \\ f & f' & g \end{pmatrix} \quad \text{and} \quad U\begin{pmatrix} a & b & e \\ c & d-1 & e'-1 \\ f & f' & g \end{pmatrix}.$$

Coefficients of the relation which come from the W coefficients with one variable equal to $1/2$, are simple algebraic functions. In a similar way, we can have a relation betwen the following nine U coefficients:

$$U\begin{pmatrix} a & b & e \\ c & d & e' \\ f & f' & g \end{pmatrix}, \quad U\begin{pmatrix} a & b & e \\ c & d & e'\pm 1 \\ f' & f' & g \end{pmatrix}, \quad U\begin{pmatrix} a & b & e \\ c & d\pm 1 & e' \\ f & f' & g \end{pmatrix} \quad \text{and} \quad U\begin{pmatrix} a & b & e \\ c & d\pm 1 & e'\pm 1 \\ f & f' & g \end{pmatrix}.$$

Furthermore, it is easily shown that

$$(ab(e)c(e')dg \mid ad(f)c(f')eg)$$

$$= (-1)^{e-f-e'+f'} \left[(2e+1)(2e'+1)(2f+1)(2f'+1) \right]^{1/2} U\begin{pmatrix} a & c & b \\ f & c & f' \\ d & e' & g \end{pmatrix}. \quad \text{(A.4)}$$

Using the definition (10) and (A, 4), we can find

$$U\begin{pmatrix} a & b & e \\ c & d & e' \\ f & f' & g \end{pmatrix} = (2g+1)\sum_{\lambda,\mu,\nu}(2\lambda+1)(2\mu+1)(2\nu+1)$$

$$\cdot U\begin{pmatrix} a & b & e \\ c & d & \lambda \\ f & \mu & \nu \end{pmatrix} U\begin{pmatrix} \lambda & \nu & e \\ c & d & e' \\ \bar{d} & \bar{h} & g \end{pmatrix} U\begin{pmatrix} \mu & b & \bar{d} \\ \nu & d & \bar{h} \\ f & f' & g \end{pmatrix}, \quad \text{(A. 5)}$$

and

$$U\begin{pmatrix} a & b & e \\ c & d & e' \\ f & f' & g \end{pmatrix} U\begin{pmatrix} \bar{a} & \bar{b} & \bar{e} \\ \bar{c} & \bar{d} & \bar{e}' \\ f & f' & g \end{pmatrix} = \sum_{\gamma_1,\gamma_2,\delta,\varepsilon}(-1)^{e'+\bar{a}-f'-\delta}(2\gamma_1+1)(2\gamma_2+1)(2\delta+1)(2\varepsilon+1)$$

$$\cdot U\begin{pmatrix} a & b & e \\ c & \varepsilon & \delta \\ f & \bar{a} & \bar{c} \end{pmatrix} U\begin{pmatrix} \bar{a} & \bar{b} & \bar{e} \\ \varepsilon & \bar{d} & \gamma_2 \\ b & f' & d \end{pmatrix} U\begin{pmatrix} c & d & e' \\ \varepsilon & \gamma_2 & \bar{d} \\ \delta & \bar{c} & \gamma_1 \end{pmatrix} U\begin{pmatrix} \bar{c} & \bar{d} & \bar{e}' \\ e & c' & g \\ \delta & \bar{\gamma}_1 & \bar{c} \end{pmatrix}. \quad \text{(A6)}$$

References

1) G. Racah, Phys. Rev. **62** (1942), 438 and **63** (1943), 367. These are referred to as RII and RIII, respectively. Also "Lectures on Group Theory: and Spectroscopy" Inst. Adv. Study, Princeton, 1951.

2) H. A. Jahn, Proc. Roy. Soc. **A201** (1950), 516 ; **205** (1951), 192 and H. A. Jahn and H. van Wieringen, ibid **A209** (1951), 502.

3) B. H. Flowers, Proc. Roy. Soc. **A210** (2951), 498 ; **212** (1952) 248 ; **215** (1951), 398 and A. R. Edmond and B. H. Flowers, ibid **A214** (1952), 515 ; **215** (1952), 398.

4) J. M. Blatt and L. C. Biedenharn, Revs. Mod. Phys. **24** (1952), 258.

5) L. C. Biedenharn and M. E. Roes, Revs. Mod. Phys. **25** (1953), 729.

6) Biedenharn, Blatt and Rose, Revs. Mod. Phys. **24** (1952), 249, referred to as BBR.

7) I. Talmi, Phys. Rev. **89** (1953), 1065.

8) H. Horie, Prog. Theor. Phys. **10** (1953), 296.

9) F. Innes, Phys. Rev. **91** (1953), 31.

10) A. Moszkowski, Phys. Rev. **89** (1953), 474.

11) A. Simon and T. A. Welton, Phys. Rev. **90** (1953), 1036.

12) L. C. Biedenharn, ORNL, No. 9098 (1952).

13) Obi, Ishidzu, Horie, Yanagawa, Tanabe and Sato, Ann. Tokyo Astr. Obs. Second Series, III, No. 3 (1953), 89.

On the Green-functions of Many-electron Problem

Shinzo NAKAI

Department of Physics, Osaka University

The Dirac-field Green-functions for many-body problem are derived with their fundamental equations from the standpoint of the S matrix theory. Further, the possibility of removing divergences contained in them together with those contained in the quantities of one-body problem is demonstrated by a consistent use of the concepts of mass and charge renormalization. On that occasion, we assume the convergence of the perturbation series currently in use after renormalization has been carried out. As our argument is, therefore, essentially based upon the assumption of expandability in powers of the coupling constant, it runs counter to Schwinger's intention. However, the physical meaning of several quantities is understood more easily in our theory.

§ 1. Introduction and summary

Many attempts have been so far made to investigate the existence of bound states when two Fermi-particles are interacting with each other through the quantized fields. First, we have a standard method[1] in which the effective potential between two particles is calculated and by using this, the two-particle Dirac (or Schroedinger) equation is solved. There is another method used by Tamm-Dancoff,[2] according to which the Schroedinger equation for the state vector has been solved with the requirement that it contains no particle-pairs and only one field quantum. These methods, however, are quite unsatisfactory; the former one can succeed only when the major effects of interaction are limited to non-relativistic region, and the latter is not fitted for the removal of divergences for the lack of covariance, as it is.

Bethe and Salpeter[3] have found the differential equation for the two-particle 'wave-function' starting from Feynman's two-body kernel. Gell-Mann and Low[4] have derived the same equation from the ordinary field-theory. Schwinger[5] has also constructed two-particle Green-function and its fundamental equation in his own new formalism which is not founded upon the expandability in powers of the coupling constant. As these equations are all relativistically covariant, we can imagine that the elimination of divergences may be dealt with by means of the usual mass and charge renormalization technique developed by Dyson.[6] However, it has not yet been verified in any of these equations. Bethe and Salpeter applied their equation practically to the ground state of deuteron based upon the scalar meson theory and solved it by iteration method in the so-called 'ladder-approximation'. But when the coupling constant is not small, e.g., when we deal with the deuteron using the pseudoscalar meson theory, it seems that we have to take up more refined interaction with the improved approximation. In fact, the existence of stationary

states of the neutron-proton system has been investigated by Lévy[7] on the basis of a two-particle Dirac equation including a retarded pseudo-scalar interaction. This interaction is treated according to two different models : in the first one, the negative energy states of each nucleon are supposed to be filled, and in the second, they are considered as empty. And he has shown that in both cases, no stable stationary state exists. If we try to reach the higher approximation[7] and to allow for the radiative corrections in that case, it is by all means necessary to establish the renormalization program for the removal of divergences. We intend in this paper to establish the renormalization scheme pertaining to two-body and, generally, many-body problems on the standpoint of Schwinger's Green-function. It is, of course, possible that we consider the same problem on the 'wave-function', for two-particle 'wave-function' is derived from the Green-function by following Gell-Mann and Low's prescription.

On the other hand, the present state of the renormalilation theory is such that, when S-matrix is expanded into the power series of the coupling constant,

$$S = \sum_{n=0}^{\infty} e^n S_n , \qquad (1 \cdot 1)$$

each term S_n has been made free from divergences owing to Dyson's elegant technique, and further that the question whether the sum of the series $(1 \cdot 1)$ does converge or not, has been discussed by several authors.[8],[9],[10] Schwinger[5] has independently developed his formal theory of Green-functions without founding on the restricted basis provided by the expandability in powers of coupling constant, though perturbation theory may be resorted to for detailed calculations. He has not, however, referred to the renormalization procedure in his paper. So noticing this point, Utiyama et al.[11] have reformulated this theory of Green-function on the side of S-matrix theory and shown that it is inevitable to expand Green-functions into the powers of the coupling constant so far as the present renormalization method is utilized.

Our renormalization procedure described below follows the one[6] in S-matrix theory, and we must, therefore, admit the series expansion in powers of the coupling constant, and moreover assume the convergence of the series. It, therefore, appears that our method of renormalization is not progressive but regressive so far as it submits to series expansions contrary to Schwinger's original standpoint, and that the renormalization independent of the expandability in powers of coupling constant ought to be developed. Electromagnetic field and electron field are here adopted as an example, simply with a view to complete understanding of Schwinger's theory in connection with S-matrix theory.

§ 2.   The derivation of Green-function for two-electron problem

As the Green-function for two-electron problem and the fundamental equation which it satisfies are derived in close connection with the S-matrix theory, the well-known theory by Tomonaga-Schwinger-Feynman-Dyson is utilized.

We take

$$L = -(1/4) \cdot F_{\mu\nu} F_{\mu\nu} - (1/2) \cdot (\partial A_\mu / \partial x_\mu)^2 - \overline{\psi}(-i\gamma_\mu \partial_\mu + m)\psi + A_\mu(j_\mu + J_\mu),$$

for the Lagrangian density and Schroedinger equation in interaction representation is written as

$$i \delta \Psi / \delta \sigma = H \Psi, \qquad H = -(j_\mu + J_\mu) A_\mu.$$

Here J_μ stands for the external current, $\bar{\psi} = \psi^* \gamma_0$ ($\psi^* = $ complex conjugate of ψ),

$$j_\mu = (e/2) \gamma^\mu_{\alpha\beta} [\bar{\psi}_\alpha, \psi_\beta] \quad \text{and} \quad F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu.$$

γ's are usual Dirac's γ-matrices which satisfy the relations,

$$\{\gamma_\mu, \gamma_\nu\} = -2\delta_{\mu\nu}, \qquad \gamma_0 = i\gamma_4, \qquad \gamma_0^2 = 1.$$

As is well-known, S-matrix is expressed in terms of Wick's[12] chronological operator T,

$$S = \sum_{n=0}^{\infty} (-i)^n / n! \cdot \int_{-\infty}^{\infty} d^4 x_1 \cdots \int_{-\infty}^{\infty} d^4 x_n T(H(x_1) \cdots H(x_n)). \tag{2.1}$$

The Green-function for electron and that for photon are given by[11]

$$G(x, x') = i(00 | T(\psi_\alpha(x), \bar{\psi}_\beta(x') S) | 00)_c = i \langle \psi_\alpha(x) \bar{\psi}_\beta(x') \rangle, \tag{2.2}$$

and

$$\mathfrak{T}_{\mu\nu}(x, x') = i(00 | T(A_\mu(x), A_\nu(x') S) | 00)'_c = i \langle A_\mu(x), A_\nu(x') \rangle', \tag{2.3}$$

respectively. Two 'zero' symbols in the bracket show that no electron and no photon are present in the initial or the final state. The notation $(\cdots T(\cdots) \cdots)'_c$ means that in expanding the T-product in the bracket into the sum of normal constituents, we retain only those terms corresponding to the peculiar Feynman-graphs. In these, all the points $x_1, x_2 \cdots x_m$, are connected with the particular points $x, x' \cdots$, where $x_1, x_2 \cdots x_n$ are the arguments of the term $T(H(x_1) \cdots H(x_n))$ included in the expression S and $x, x' \cdots$ are those of the operators contained in the bracket. Acordingly, all the points $x_1, x_2 \cdots x_n, x, x' \cdots$ in the expression $(\cdots T(\cdots) \cdots)'_c$ are mutually connected, whereas in $(\cdots T(\cdots) \cdots)_c$ they may be divided into several sets of points and all the points in each set are mutually connected, but any pair of points separately belonging to different sets is disconnected. It is to be remarked, in passing, that regarding the angular brackets, the following relation

$$\langle A(x), B(x') \rangle = \langle A(x), B(x') \rangle' + \langle A(x) \rangle \langle B(x') \rangle$$

holds.

If we regard the space-time coordinates as matrix indices, the fundamental equations are given by*

$$\{\gamma(P - e\langle A \rangle) + m + \Sigma^*\} G = 1, \qquad (\Box \delta_{\mu\rho} - P_{\mu\rho}) \mathfrak{T}_{\rho\nu} = -\delta_{\mu\nu}. \tag{2.4}$$

Lastly, as the relation,

$$ie\gamma_\mu \delta G(x, x') / \delta J_\mu(x) = \int \Sigma^*(x, y) G(y, x') d^4 y, \tag{2.5}$$

---

* For the definitions of Σ^*, $P_{\mu\rho}$ etc., see the appendix 1.

is confirmed, all the equations given by Schwinger[5] can be obtained.

The Dirac field Green-function G_{12} for two-body problem is explicitly given by

$$G_{12}(x_1 x_2, x_1', x_2') = \langle \psi(x_1)\psi(x_2)\bar{\psi}(x_1')\bar{\psi}(x_2') \rangle$$

$$= \sum_{n=0}^{\infty} (-i)^n/n! \cdot \int d^4 y_1 \cdots \int d^4 y_n (0| T(\psi(x_1)\psi(x_2)\bar{\psi}(x_1')\bar{\psi}(x_2') \cdot H(y_1) \cdots H(y_n))|0)_c.$$

$$(2\cdot6)$$

As the definition $(2\cdot6)$ shows, any diagram of G_{12} has a pair of electron-lines (λ_1 and λ_2) which terminates at x_1 and x_2, and another pair of electron-lines (λ_1' and λ_2') which leaves x_1' and x_2', inclusive of such a case as $\lambda_1 = \lambda_1'$ and $\lambda_2 = \lambda_2'$. Since $\langle \psi \rangle = \langle \bar{\psi} \rangle = 0$ owing to the conservation of the electric charge, it is assured that diagram of G_{12}, as a whole, consists of two open polygons and several cloosed loops of electron-lines. And it is also assured that they are connected with each other by photon-lines. (See fig. 1.) Now we draw the directed line (electron-line) joining y' to y for every internal line which arises from the Wick's[12] contraction procedure,

$$(\psi_\alpha(y)\bar{\psi}_\beta(y')) = T(\psi_\alpha(y)\bar{\psi}_\beta(y')) - N(\psi_\alpha(y)\bar{\psi}_\beta(y')) = S_{\alpha\beta}^F(y-y').$$

Then, it is supposed that two electron-open-polygons form two trees, which may be connected with each other by photon lines. Regarding the vertices x_1' and x_2' as the roots of the trees, a unique 'up-direction' is fixed in every directed line, namely, the direction pointing away from x_1' and x_2'. Let us analyse $(2\cdot6)$ 'downwards' from the top-vertices x_1 and x_2. For convenience, we use a term 'a graph of a T-product', by which is meant a Feynman-graph corresponding to any one of the normal constituents arising from the Wick's expansion of the T-product. As already stated, all the points $y_1 \cdots y_n$ in every allowable graph of the T-product in $(2\cdot6)$ are connected at least either with x_1 and x_1' or with x_2 and x_2' by photon lines or electron lines. In every such graph, there are always two electron-bridge-lines*, ρ_1 and ρ_2, which are connected with x_1 and x_2, respectively. (fig. 1.) The vertices $y_1 \cdots y_n$ may be divided into three sets of vertices, a_1, a_2 and b, by these bridge-lines as follows.
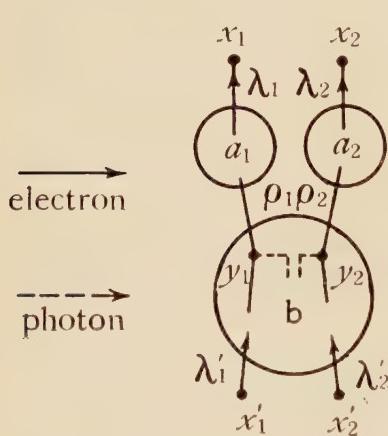
Fig. 1.

1) Every vertex in set a_1, if connected with x_2, cannot be connected without passing through ρ_1 and ρ_2.**

2) The vertices contained in set b can be connected neither with x_1 without passing through the bridge-line ρ_1 nor with x_2 without crossing the bridge-line ρ_2.

---

* By this word, is meant the electron-line by which any Feynman-graph is divided into two separate parts.

** This fact expresses at the same time that every vertex in set a_2, if connected with x_1, cannot be connected without passing through ρ_1 and ρ_2.

Then, there arise the following two cases.

1st; The set b is empty. Hence, the bridges, ρ_1 and ρ_2, happen to coincide with λ_1' and λ_2', respectively. In this case, (2·6) turns into

$$G(x_1 x_1') G(x_2 x_2') - G(x_1 x_2') G(x^2 x_1').$$

2nd; The set b is not empty. In this case, there may be, in general, several ways to chose a bride-line in accordance with the situations as to which line is chosen for a bridge-line on each electron-open-polygon in any allowable graph. Hence, let us select a special bridge-line so as to make both of a_1 and a_2 the maximum set on each electron-open-polygon. (2·6) may be transformed into

$$\int G(x_1 y_1) G(x_2 y_2) e^2 \langle \gamma A \phi(y_1), \gamma A \phi(y_2), \bar{\phi}(x_1'), \bar{\phi}(x_2') \rangle'' d^4 y_1 \, d^4 y_2$$

$$= \int G(x_1 y_1) G(x_2 y_2) W_{12}(y_1 y_2, x_1' x_2') d^4 y_1 d^4 y_2,$$

where $\langle \cdots \rangle''$ stands for such peculiar diagrams out of the ones belonging to $\langle \cdots \rangle'$ that both vertices y_1 and y_2 are included in the 'proper-connected' graph manifesting two-body interaction.* Let us say that the vertices, y_1 and y_2, have the 'quality-P' in order to express this quality briefly.

Consequently, we can find

$$G_{12}(x_1 x_2, \ x_1' x_2') = G(x_1 x_1') G(x_2 x_2') - G(x_1 x_2') G(x_2 x_1')$$

$$+ \int G(x_1 y_1) G(x_2 y_2) W_{12}(y_1 y_2 x_1' x_2') d^4 y_1 d^4 y_2. \qquad (2\cdot7)$$

When this equation is multiplied by $(\gamma \pi + M)_1 (\gamma \pi + M)_2$ from the left, where

$$\pi(x) = P - e \langle A(x) \rangle, \qquad M(x, y) = m \delta(x - y) + \Sigma^*(x, y),$$

then by using (2·4) we obtain

$$(\gamma \pi + M)_1 (\gamma \pi + M)_2 G_{12} - W_{12} = 1_{12}$$

$$(2\cdot8)$$

where

$$W_{12}(x_1 x_2, \ x_1' x_2') = \{I_{12} G_{12}\}(x_1 x_2, \ x_1' x_2') = e^2 \langle \gamma A \phi(x_1), \gamma A \phi(x_2), \bar{\phi}(x_1'), \bar{\phi}(x_2') \rangle'',$$

$$1_{12} = \delta(x_1 - x_1') \delta(x_2 - x_2') - \delta(x_1 - x_2') \delta(x_2 - x_1').$$

It is also understood that if the analysis of (2·6) is repeated 'upwards' from the vertices, x_1' and x_2', we get the adjoint equation

$$G_{12}(\gamma \pi + M)_1 (\gamma \pi + M)_2 - W_{12}' = 1_{12}, \qquad (2\cdot9)$$

---

* Here we use the term 'a proper-connected diagram in two-body problem', by which is meant such a diagram that is consisting of two electron-open-polygons and closed loops connected by photon lines, and can be no more separable by two electron-bridge-lines into two parts. But in the special case when the diagram includes no internal electron-lines and only one vertex on any particular open polygon, the bridge-line, only on this polygon, is to be replaced by the very vertex.

where

$$W_{12}' = G_{12}I_{12} = e^2 \langle \psi(x_1),\ \psi(x_2),\ \overline{\psi}\gamma A(x_1'),\ \overline{\psi}\gamma A(x_2')\rangle''. \qquad (\text{ fig. 2.})$$

These equations are those given by Schwinger. We must, however, complete our analysis of the interaction term I_{12} between two electrons from the present standpoint before we compare our results with those of Schwinger's.[5]

Next let us proceed to the analysis of I_{12}. For this purpose, we had better deal with the expression $I_{12}G_{12}$ first.

$$\{I_{12}G_{12}\}\ (y_1 y_2,\ x_1'x_2') = \sum_{n=0}^{\infty}((-i)^n/n\,!)\cdot e^2 \int d^4z_1 \cdots \int d^4z_n (0\,|\,T(\gamma A\psi(y_1),$$

$$\gamma A\psi(y_2),\ \overline{\psi}(x_1'),\ \overline{\psi}(x_2'),\ H(z_1)\cdots H(z_n)\,|0)''. \qquad (2\cdot10)$$
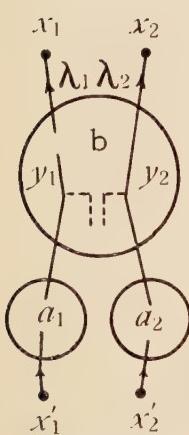
Suppose now that two electron-bridge-lines $(\rho_1',\ \rho_2')$ are chosen out of any one of allowable graphs of the T-product in $(2\cdot10)$, where n is fixed. (fig. 3) Each open polygon of the electron lines necessarily contains only one bridge-line. Then the vertices of this diagram are divided into two parts α and β. The points in β cannot be connected with x_1' and x_2' without crossing the bridge-lines ρ_1' and ρ_2'. This time, β cannot be empty, for photon lines are linked to both vertices, y_1 and y_2. Since we may generally have several ways to choose a bridge-line on each electron polygon, particular bridge-lines are taken so that α may be the maximum set. Accordingly, β is no more separable by two bridge-lines into two parts, that is, any diagram of β is a 'proper-connected' diagram in the two-body problem. When we thus divide any diagram of $(2\cdot10)$ into two parts, there are four cases corresponding to*

Fig. 2.

Fig. 3.

$$\begin{array}{llll}
\text{I)} & \rho_1'=\lambda_1'',\ \rho_2'=\lambda_2'' & \text{III)} & \rho_1'=\lambda_1'',\ \rho_2'\not\equiv\lambda_2'' \\[4pt]
\text{II)} & \rho_1'\not\equiv\lambda_1'',\ \rho_2'=\lambda_2'' & \text{IV)} & \rho_1'\not\equiv\lambda_1'',\ \rho_2'\not\equiv\lambda_2''
\end{array} \Bigg\} \cdot(2\cdot11)$$

Repeating similar considerations about all the allowable constituents appearing in $(2\cdot10)$ and summing up over all the contributions from the sets α's and β's, we can finally write down $(2\cdot10)$ in each case as follows.

I)

$$e^2\gamma_1{}^\mu\gamma_2{}^\nu\langle\psi(y_1)\psi(y_2)\overline{\psi}(x_1')\overline{\psi}(x_2')\rangle\langle A_\mu(y_1)A_\nu(y_2)\rangle'$$

$$= -i\,e^2\gamma_1{}^\mu\mathfrak{D}_{\mu\nu}(y_1-y_2)\gamma_2{}^\nu G_{12}(y_1 y_2,\ x_1'x_2') = U_I G_{12}\ .$$

(fig. 4a)

* λ_1'' and λ_2'' are the electron-lines which are directly connected with the vertex y_1 and the vertex y_2, respectively.

II)

$$ie^2 \int \langle \gamma A(y_1), \gamma A\psi(y_2), \bar{\psi}e\gamma A(y_2')\rangle'_{P.C.}\langle \psi(y_1)\psi(y_2')\bar{\psi}(x_1')\bar{\psi}(x_2')\rangle d^4y_2' \;.$$

If one notes here that $\langle \cdots \rangle'_{P.C.}$ includes only those diagrams which are 'proper' and 'connected', $A(y_1)$ cannot be directly paired with $A(y_2)$ or $A(y_2')$. Accordingly, one chooses $A_\nu(\hat{\xi}_2)$ for the partner of $A(y_1)$, where the vertex $\hat{\xi}_2$ is situated between the vertices, y_2 and y_2', and is 'proper-connected' with both of them. Then, the above formula is transformed into

$$= -ie^2\gamma_1^\mu \int \langle A_\mu(y_1), A_\nu(\hat{\xi}_2)\rangle' i\langle \gamma A\psi(y_2), J_\nu(\hat{\xi}_2), \bar{\psi}e\gamma A(y_2')\rangle'.$$

$$G_{12}(y_1 y_2', \; x_1' x_2') d^4 y_2' d^4\hat{\xi}_2$$

$$= -ie^2\gamma_1^\mu \int \mathfrak{D}_{\mu\nu}(y_1 - \hat{\xi}_2)\Lambda^\nu(y_2 y_2', \hat{\xi}_2) G_{12}(y_1 y_2', \; x_1' x_2') d^4 y_2' d^4\hat{\xi}_2 = U_{II} G_{12}. \quad \text{(fig. 4b)}$$

III)   In the same way as in II), $A_\nu(\hat{\xi}_1)$ is taken as a partner of $A(y_2)$, where the vertex $\hat{\xi}_1$ is situated between the vertices, y_1 and y_1' and is proper-connected with both of them.   Therefore, we obtain

$$= -ie^2 \int \Lambda^\mu(y_1 y_1', \hat{\xi}_1)\mathfrak{D}_{\mu\nu}(\hat{\xi}_1 - y_2)\gamma_2^\nu G_{12}(y_1' y_2, \; x_1' x_2') d^4\hat{\xi}, \; d^4 y_1'$$

$$= U_{III} G_{12} \;. \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{(fig. 4c)}$$

IV)   In this case, we get from $(2\cdot10)$

$$e^2 \int d^4 y_1' d^4 y_2' \langle \gamma A\psi(y_1), \gamma A\psi(y_2), \bar{\psi}e\gamma A(y_1'), \bar{\psi}e\gamma A(y_2')\rangle'_{P.C.} \; G(y_1' y_2', \; x_1' x_2')$$

$$= U_{IV} G_{12} \;. \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{(fig. 4d)}$$
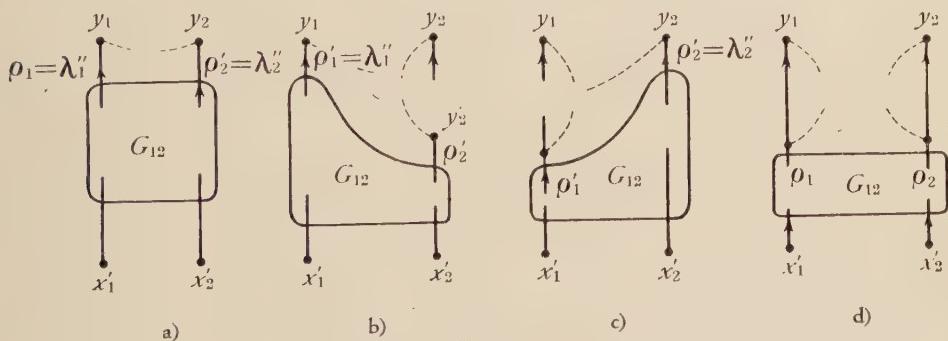
Fig. 4

Thus $(2\cdot10)$ finally turns into

$$I_{12} G_{12} = (U_I + U_{II} + U_{III} + U_{IV}) G_{12}, \qquad\qquad (2\cdot12)$$

from which the explicit expression of I_{12} can be obtained. U_{IV}, however, cannot be written down simply in terms of the quantities G, I', $\mathfrak{D}$ etc., used in one-body problem.

We must, therefore, devise the means to find the recurrence formula for I_{12}. It is the shorter way to deal with the expression $(U_{III}+U_{IV})G_{12}$. Now notice one internal photon line in any diagram connecting the vertex z_1 with the vertex z_2. Next we suppose the graph which is obtained from the above diagram by removing the very internal line and by putting the external potentiol $\langle A(z_2)\rangle$ on the very vertex z_2. Then, the relation,

$$\mathfrak{I}_{\mu\nu}(z_1-z_2)=\delta\langle A_\mu(z_2)\rangle/\delta J_\nu(z_1)=\delta/\delta J_\nu(z_1)\cdot\int \mathfrak{I}_{\mu\rho}(z_1-z_2)J_\rho(z_3)d^4z_3,$$

enables us to use the latter diagram instead of the former one with the addition of the operation $\delta/\delta J$. And comparing these two diagrams, we find that the number of internal photon lines is different by one in them. So our purpose of deriving the recurrence formula for I_{12} is attained as follows. It is clear by the above analysis that both vertices, y_1 and y_2, are belonging to the proper-connected graph manifesting two-body interaction, that is, the vertices, y_1 and y_2, have the 'quality-P'. Now when the photon line linked to the vertex y_1 is removed, there may be considered the following two cases about the vertices, y_1 and y_2.

Fig. 5.

i) Neither y_1 nor y_2 has the 'quality-P'.
ii) Vertex y_1 has not the 'quality-P', but vertex y_2 has still the 'quality-P'.
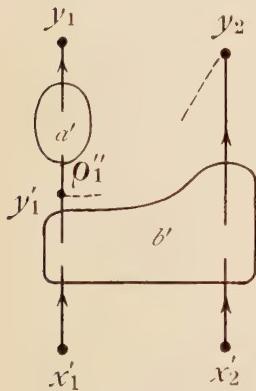
And i) happens in the cases, I) and II) at $(2\cdot11)$, and ii) happens in the cases, III) and IV). As we are now concerned with only $(U_{III}+U_{IV})G_{12}$, we have only to treat the case ii). We analyse $(2\cdot10)$ 'down-wards' from the vertex y_1 and divide it into two pieces. (fig. 5) The division is made in such a way that the vertices contained in b' cannot be connected with y_1 without crossing the bridge-line. On choosing the bridge-line in any allowable diagram, we take a particular one (ρ_1'') which makes the set a' maximum. Then the farther end-vertex of ρ_1'' from y_1, which we denote by y_1', has the 'quality-P'. Accordingly, we can find from $(2\cdot10)$

$$(U_{III}+U_{IV})G_{12}=\int \gamma_1^{\rho}G(y_1-y_1')e^a\langle A_\rho(y_1),\gamma A\psi(y_1'),\gamma A\psi(y_2),\bar\psi(x_1'),\bar\psi(x_2')\rangle''d^4y_1'$$

$$=-ie\gamma_1^{\rho}\int G(y_1-y_1')\cdot\delta/\delta J_\rho(y_1)\cdot e^2\langle\gamma A\psi(y_1'),\gamma A\psi(y_2),\bar\psi(x_1'),\bar\psi(x_2')\rangle''$$

$$=-ieT_\rho[\gamma_1 G_1\,\delta/\delta J](I_{12}G_{12}). \tag{2·13}$$

Consequently, $(2\cdot12)$ becomes

$$I_{12}G_{12}=-ie^2T_\rho[\gamma_1 I_{12}'\mathfrak{I}]-ieT_\rho[\gamma_1 G_1\,\delta/\delta J](I_{12}G_{12}). \tag{2·14}$$

In the same way, it is also found that the equation,

$$I_{12}G_{12}=-ie^2T_\rho[\gamma_2 I_{12}'\mathfrak{I}]-ieT_\rho[\gamma_2 G_2\,\delta/\delta J](I_{12}G_{12}) \tag{2·15}$$

holds. These equations are the integro-differential equations for $I_{12}G_{12}$ given by Schwinger. Therefore,

$$\{(\gamma\pi + M)_1(\gamma\pi + M)_2 - I_{12}\}G_{12} = 1_{12} \qquad (2\cdot16)$$

is derived from $(2\cdot8)$ and $(2\cdot14)$.

Using the relations, $\delta G_{12}/\delta J = -(1/2)G_{12}(\delta G_{12}^{-1}/\delta J)G_{12}$ and

$$\frac{\delta}{\delta J}\{I_{12}G_{12}\} = \frac{1}{2}\left\{(1_{12} + I_{12}G_{12})\frac{\delta I_{12}}{\delta J} - I_{12}(1_{12} + G_{12}I_{12})\left(G_1\frac{\delta G_1^{-1}}{\delta J} + G_2\frac{\delta G_2^{-1}}{\delta J}\right)\right\}G_{12},$$

we can derive the expression

$$I_{12} = -ie^2 T_p[\gamma_1 I'_2\mathfrak{T}] - ie^2/2 \cdot T_p[\gamma_1 G_1(1_{12} + I_{12}G_{12})\delta/c\delta J \mid I_{12}$$

$$+ ie_2/2 \cdot T_p[\gamma_1 G_1 I_{12}(1_{12} + G_{12}I_{12})G_1 \delta/c\delta J]G_1^{-1}$$

$$+ ie^2/2 \cdot T_p[\gamma_1 G_1 I_{12}(1_{12} + G_{12})G_2 \delta/e\delta J]G_2^{-1} \qquad (2\cdot17)$$

for I_{12}. We can also carry out the similar calculation about $(2\cdot9)$.

§ 3. The formulation of the Green-function for many-body problem

We shall derive in this section the Green-function for n-electron problem and its fundamental equation when those for less than n-electron problem are already known. The Green-function for n-electron problem is defined in terms of[*]

$$G_N(x_1\cdots x_n, \ x_1'\cdots x_n') = i_n(-1)^{\overset{n-1}{\underset{m=1}{\Sigma}}}\langle\phi(x_1)\cdots\phi(x_n), \ \bar\psi(x_1')\cdots\phi(x_n')\rangle. \qquad (3\cdot1)$$

Any graph of G_N consists of the following three parts; n pieces of electron-lines, $\lambda_1\cdots\lambda_n$, which terminate at the end-points, $x_1\cdots x_n$, n pieces of electron-lines, $\lambda_1'\cdots\lambda_n'$, which leave the end-points, $x_1'\cdots x_n'$, and the part between the above stated two sets of electron-lines, which is established by the contributions of S-matrix and generally may contain the complex n-body interaction.

The analysis of $(3\cdot1)$ is made quite analogously to $(2\cdot6)$. That is, every Feynman graph given by $(3\cdot1)$ is supposed to form the figure of n trees and the vertices, $x_1'\cdots x_n'$, are regarded as the roots of the trees. The unique 'up-direction' in every directed line, then, can be determined by the direction pointing away from $x_1'\cdots x_n'$. We analyse $(3\cdot1)$ 'downwards' from the top-vertices $x_1\cdots x_n$. The vertices, $y_1\cdots y_l$, in any allowable graph of the T-product in $(3\cdot1)$ may be divided into three sets of vertices (α, β and γ) by the bridge-line-set Ω_{N-1} (constituted by $\omega_1\cdots\omega_{n-1}$) and the bridge-line ω_n in the following way. (fig. 6)

1) Every vertex in the set α, if connected with x_n, is connected necessarily by the bridge-line ω_n.

---

[*] We use $G_{i_1\cdots i_m}$, $I_{i_1\cdots i_m}$ etc., too, in place of G_M, I_M etc. to show their arguments explicitly.

2)   The vertices included in the set γ must pass through Ω_{N-1} and ω_n, respectively, to get to the top-vertices $x_1 \cdots x_{n-1}$ and x_n. Then, there are the following two cases.

1st; the set γ is empty. The bridge-lines of Ω_{N-1} and the bridge-line ω_n, therefore, happen to coincide with $\lambda_1' \cdots \lambda_{n-1}'$ and λ_n'. In this case, $(3 \cdot 1)$ becomes

$$\sum_{i=1}^{n} {}'(-1)^{n+i} G_{N-1}(1 \cdots \overline{n-1},\ 1' \cdots \overline{i-1'},\ \overline{i+1'} \cdots n') G(ni'). \qquad (3 \cdot 2)$$

For short, $x_1 \cdots x_n$ and $x_1' \cdots x_n'$ are written as $1 \cdots n$ and $1' \cdots n'$.

2nd; the set γ is not empty. In this case, though we have several ways to take a bridge-line and a bridge-line-set, we take them so as to make both α and β maximum. Now denote those vertices by $y_1 \cdots y_n$, which are the nearer end-points of each bridge-line to $x_1' \cdots x_n'$. There arise, then, the following $(n-1)$ cases from this division of the diagram, correspondingly to the situations that m vertices of $y_1 \cdots y_{n-1}$ are involved in the ' proper-connected '* diagrams, (where m runs from one to $n-1$.)   These diagrams stand for the $(m+1)$-body interaction inclusive of the vertex y_n. Here, a complication arises pertaining to the set γ except for $m=n-1$. We take m vertices out of $y_1 \cdots y_{n-1}$ and label them as $y_{i_1} \cdots y_{r_m}$. The rest are denoted by $y_{i_{m+1}} \cdots y_{i_{n-1}}$. Then all or some of $(n-m-1)$ vertices, $y_{i_{m+1}} \cdots y_{i_{n-1}}$, may agree with all or some of $(n-m-1)$ vertices out of $x_1' \cdots x_n'$. The set γ is wanting in the vertices on some open polygons. And

Fig. 6.

$$(\alpha_I + \gamma = \delta + \varepsilon)$$
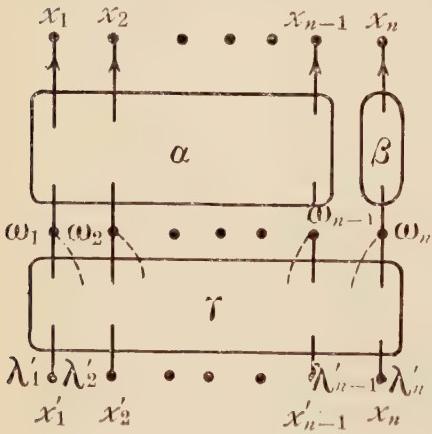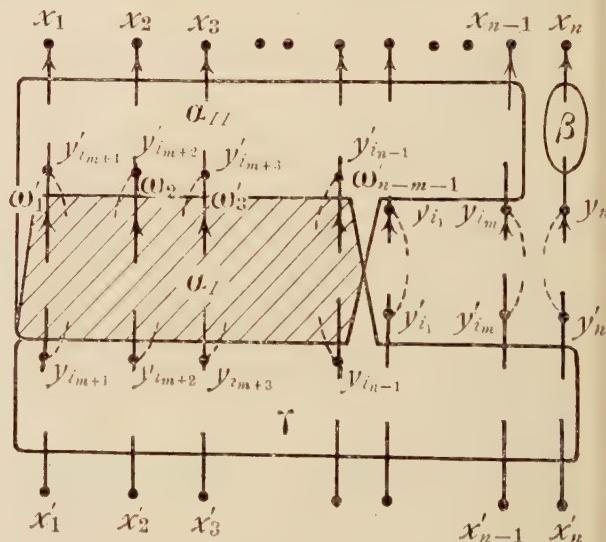
Fig. 7.

---

* A ' proper-connected ' diagram in n-body problem is such that cannot be divided by n pieces of bridge-lines into two parts. However, in the special case when the diagram includes only one vertex and no internal electron-lines on any particular electron-open-polygon, the bridge-line, only on this polygon, is to be replaced qv the very vertex.

this fact brings about the trouble as we try further to analyse the set γ. Hence, a supplemental procedure is needed in order to avoid it. So, we think of dividing the set α further into two pieces by the bridge-line-set, $\omega_1' \cdots \omega_{n-m-1}'$, and making up for the vertices wanting in γ by supplementing a part of α. The division of the set α is defined as follows. (fig, 7) Each one of $(n-m-1)$ electron-bridge-lines, $\omega_1' \cdots \omega_{n-m-1}'$, is necessarily situated separately on each of $(n-m-1)$ open polygons of the electron-lines. α is divided two parts, α_I and α_{II}. The vertices contained in the set α_{II} cannot be connected with the vertices $y_{i_{m+1}} \cdots y_{i_{n+1}}$ without crossing the bridge-lines. And if we choose the bridge-line-set $\omega_1' \cdots \omega_{n-m-1}'$ so as to make α_I maximum, the set is uniquely fixed in the set α. By the way, it is to be remarked that all or some of the vertices $y_{i_{m+1}}' \cdots y_{i_{n-1}}'^{*}$ may happen to coincide with the corresponding number of the end-points out of $x_1 \cdots x_n$. We consider the sum of the set γ and the set α_I inclusive of the bridge-lines $\omega_1' \cdots \omega_{n-m-1}'$. Then, they contain necessarily at least one vertex on any one of n electron-open-polygons. We think of dividing the vertices in $(\alpha_I + \gamma)$
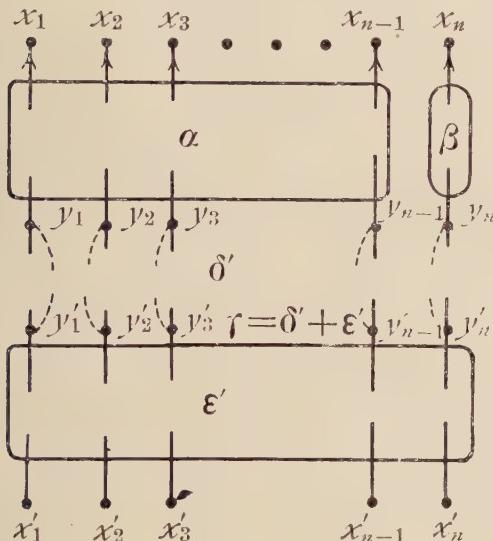
Fig. 8.

into two parts, ∂ and ε, by a new bridge-line-set consisting of n electron-lines by means of the similar arguments to the analysis of $(2 \cdot 10)$. ε is chosen so that it may be the maximum set and every point in it may be unable to be connected with $y_{i_{m+1}}' \cdots y_{i_n}'$ without crossing the bridge-line-set. ∂ then stands for the $(m+1)$-body interaction and this is assured when we recall the definition Γ_{M+1}^{**}. We make the above-stated analysis consisting of two steps about all the allowable constituents in $(3 \cdot 1)$, and sum them up in each set separately. On that occasion, if one considers the fact that the set α_I is utilized not only in the first step but also in the second, we obtain the following expressions from $(3 \cdot 1)$.

$m = 1$;

$$(1/(n-1)!) \sum_{i=1}^{n-1} \int G_{N-1}(1 \cdots \overline{n-1}, 1'' \cdots \overline{n-1}'') G(n, n'')$$

$$\times G_{N-2}^{-1}(1'' \cdots \overline{i-1}'', i+1'' \cdots \overline{n-1}'', 1''' \cdots \overline{i-1}''', i+1''' \cdots \overline{n-1}''')$$

$$\times I_{II}(i''n'', i'''n''') G_N(1''' \cdots n''', 1' \cdots n') d(1'' \cdots n'') d(1''' \cdots n'''),$$

---

* $y_{i_{m+1}}' \cdots y_{i_{n-1}}'$ are, respectively, the nearer end-points of bridge-lines $\omega_1' \cdots \omega_{n-m-1}'$ to $x_1 \cdots x_n$.
** With reference to the definition I_{M+1}, see $(2 \cdot 15)$ and $(3 \cdot 9)$.

$m = m \leq n - 2$;

$$(1/(n-1\,!))\sum_{(i_1\cdots i_m)}^{(n-1)}\int G_{N-1}(1\cdots\overline{n-1},\ 1''\cdots\overline{n-1}'')\,G(n,\ n'')$$

$$\times\, G_{N-M-1}^{-1}(1''\cdots\overline{n-1}'',\ 1'''\cdots\overline{n-1}''')\,I_{M+1}(i_1''\cdots i_m'',\ n'',\ i_1'''\cdots i_m''',\ n''')$$

$$\times\, G_N(1'''\cdots n''',\ 1'\cdots n')\,d(1''\cdots n'')\,d(1'''\cdots n''').\qquad(3\cdot3)$$

Here, $1\cdots n$ and $d(1''\cdots n'')$ are used instead of $x_1\cdots x_n$ and $dy_1''\cdots dy_n''$, respectively. Moreover, $\sum_{(i_1\cdots i_m)}^{(n-1)}$ stands for the summation over all the combinations of m vertices, $y_{i_1}''\cdots y_{i_m}''$, which are taken out of $(n-1)$ vertices, $y_1''\cdots y_{n-1}''$. The factor $1/(n-1)!$ stands in order to remove the repeated use of the same diagrams which arises from the antisymmetrical function G_{N-1}.

Lastly, in the case of $m = n-1$, the overlapping like the previous $(m \leq n-2)$ cases does not occur. (fig. 8) Accordingly, we can directly divide γ into two sets, δ' and ε' by bridge-lines. The set δ' stands for the n-body interaction and consists of $\sum_{k=0}^{n}{}_nC_k$ subsets. ${}_nC_k$ subsets of them are those in which k vertices of $y_1'\cdots y_n'$ coincide with the corresponding number of $y_1\cdots y_n$. Therefore, $(3\cdot1)$ becomes, in this case,

$$(1/(n-1)!)\int G_{N-1}(1\cdots\overline{n-1},\ 1''\cdots\overline{n-1}'')\,G(n,\ n'')\,I_N(1''\cdots n'',\ 1'''\cdots n''')$$

$$\times\, G_N(1'''\cdots n''',\ 1'\cdots n')\,d(1''\cdots n'')\,d(1'''\cdots n''')\qquad(3\cdot4)$$

where

$$\int I_N(y_1\cdots y_n,\ y_1'\cdots y_n')\,G_N(y_1'\cdots y_n',\ x_1'\cdots x_n')\,dy_1'\cdots dy_n'$$

$$= i^n(-)^{\overset{n-1}{\underset{}{\sum}}j}\langle e_\Upsilon A\psi(y_1)\cdots e_\Upsilon A\psi(y_n),\ \bar\psi(x_1')\cdots\bar\psi(x_n')\rangle''.\qquad(3\cdot5)^*$$

From $(3\cdot2)-(3\cdot4)$, we obtain

$$G_N(1\cdots n,\ 1'\cdots n') = \sum_{j=1}^{n}(-1)^{n+j}G_{N-1}(1\cdots\overline{n-1},\ 1'\cdots\overline{j-1}',\ \overline{j+1}'\cdots n')\,G(nj')$$

$$+\sum_{m=1}^{n-2}\sum_{(i_1\cdots i_m)}^{(n-1)}\frac{1}{(n-1)!}\int G_{N-1}(1\cdots\overline{n-1},\ 1''\cdots\overline{n-1}'')\,G(n,\ n'')$$

$$\times\, G_{N-M-1}^{-1}(\underbrace{1''\cdots\overline{n-1}''}_{},\ \underbrace{1'''\cdots\overline{n-1}'''}_{})\,I_{M+1}(i_1''\cdots i_m'',\ n'',\ i_1'''\cdots i_m''',\ n''')$$

$$(n-m-1)\ \text{factors except}\ (i_1\cdots i_m)$$

$$\times\, G_N(1'''\cdots n''',\ 1'\cdots n')\,d(1''\cdots n'')\,d(1'''\cdots n''')$$

$$+\int\frac{1}{(n-1)!}\,G_{N-1}(1\cdots\overline{n-1},\ 1''\cdots\overline{n-1}'')\,G_n(n,\ n'')\,I_N(1''\cdots n'',\ 1'''\cdots n''')$$

$$\times\, G_N(1'''\cdots n''',\ 1'\cdots n')\,d(1''\cdots n'')\,d(1'''\cdots n''').\qquad(3\cdot6)$$

---

* The expression $\langle\ldots\rangle''$ stands for the peculiar diagrams out of those belonging to $\langle\ldots\rangle'$. In them, all of vertices, $y_1\cdots y_n$, have the special quality expressed in terms of the 'quality-P'. (See p. 159)

When $G_{12\cdots\overline{n-1}}^{-1}G_n^{-1}$ is multiplied from the left, $(3\cdot6)$ turns into

$$G_{12\cdots n}^{-1}\,G_{12\cdots n}=1_{12\cdots n}\,,\qquad\qquad (3\cdot7)$$

where

$$1_{12\cdots n}=\sum_{p(i_1\cdots i_n)}\hat o(x_1-x_{i_1})\hat o(x_2-x_{i_2})\cdots\hat o(x_n-x_{i_n})\epsilon(i_1\cdots i_n).$$

$\displaystyle\sum_{p(i_1\cdots i_m)}$ means the summation of all the permutation of $12\cdots n$, and ϵ is given by

$$\epsilon(i_1\cdots i_n)=\begin{cases}+1,&\text{if }i_1\cdots i_n\text{ is even permutation of }1\cdots n,\\[2mm]-1,&\text{if }i_1\cdots i_n\text{ is odd permutation of }1\cdots n.\end{cases}$$

Further, when we use the abbreviations, $G_n^{-1}=(\gamma\pi+M)_n$, $\pi=P-e\langle A\rangle$ and $M=m+\sum^*$,

$$G_{12\cdots n}^{-1}=G_{12\cdots\overline{n-1}}^{-1}\,G_n^{-1}-\Big\{\sum_{m=1}^{n-2}\sum_{(i_1\cdots i_m)}^{n-1}\underbrace{G_{12\cdots\overline{n-1}}^{-1}}_{(n-m-1)\text{ factors except }(i_1\cdots i_m)}I_{i_1\cdots i_{m,n}}+I_{12\cdots n}\Big\}.\qquad (3\cdot8)$$

For example,

$$G_{123}^{-1}=G_{12}^{-1}G_3^{-1}-(G_1^{-1}I_{23}+G_2^{-1}I_{31}+I_{123})$$

$$=G_1^{-1}G_2^{-1}G_3^{-1}-(G_1^{-1}I_{23}+G_2^{-1}I_{31}+G_3^{-1}I_{12}+I_{123})\qquad(\because G_{12}^{-1}=G_1^{-1}G_2^{-1}-I_{12})$$

$$G_{1234}^{-1}=G_{123}^{-1}G_4^{-1}-\Big\{\sum_{(i_1i_2)}^3 G_{i_1i_2}^{-1}I_{i_3 4}+\sum_{(i_1)}^3 G_{i_1}^{-1}I_{i_2 i_3 4}+I_{1234}\Big\}$$

$$=G_1^{-1}G_2^{-1}G_3^{-1}G_4^{-1}-\Big\{\sum_{(i_1i_2)}^4 G_{i_1}^{-1}G_{i_2}^{-1}I_{i_3 4}+\sum_{(i_1)}^3 G_{i_1}^{-1}G_4^{-1}I_{i_2 i_3}+\sum_{(i_1)}^4 G_{i_1}^{-1}I_{i_2 i_3 4}$$

$$+G_4^{-1}I_{123}+I_{1234}\Big\}+(1/2)\sum_{(i_1i_2)}^4 I_{i_1 i_2}I_{i_3 i_4},\quad\text{etc.}$$

The expression for $I_{12\cdots n}$ can be obtained if we think of dividing $(3\cdot5)$ into two parts, i.e., $I_{12\cdots n}$ and $G_{12\cdots n}$, by a bridge-line-set consisting of n electron-lines. The prescription for this division is quite analogous to that for the analysis of I_{12}. We shall write the results only.

i)
$$I_{12\cdots n}=\sum_{m=0}^n I_{12\cdots n}(m)$$

ii)
$$I_{12\cdots n}(n)=e^n\langle\gamma A(1),\ \gamma A(2),\ \cdots,\ \gamma A(n)\rangle'_{P.C.}\,,$$

$$n-1\ge m\ge1,\ \ (m=n-k),\qquad\qquad k\ \text{factors}$$

$$I_{12\cdots n}(m)=i^k\cdot e^{n-k}\sum_{(i_1\cdots i_k\ne n)}'\langle\gamma A(1\cdots\gamma A(n),\ e\gamma A\psi(i_1)\overline{\psi}e\gamma A(i_1')\cdots e\gamma A\psi(i_k)\overline{\psi}e\gamma A(i_k')\rangle'_{P.C.}$$
$$\underbrace{\qquad\qquad}_{(n-k)\text{ factors except }(i_1\cdots i_k)}$$

$$+i^k e^{n-k}\sum_{(i_1\cdots i_k=n)}''\langle\gamma A(1)\cdots\gamma A(n),\ e\gamma A(i_1)\ \overline{\psi}e\gamma A(i_1')\cdots e\gamma A\phi(i_k)\overline{\psi}e\gamma A(i_k')\rangle'_{P.C.}$$
$$\underbrace{\qquad\qquad\qquad}_{(n-k)\text{ factors except }(i_1\cdots i_k)}\qquad\underbrace{\qquad\qquad}_{k\text{ factors}}$$

iii) $\ I_{12\cdots n}(0)=i^n\langle e\gamma A\phi(1),\ \overline{\psi}e\gamma A(1'),\cdots e\gamma A\phi(n),\ \overline{\psi}e\gamma A(n')\rangle'_{P.C.}.\qquad (3\cdot9)$

The sum of i) and the first term of ii) can be rewritten as

$$-i e \, T_\rho \left[\gamma_n \partial/\partial J \right] I_{12 \ldots \overline{n-1}}.$$

And we can obtain

$$I_{12 \ldots n} G_{12 \ldots n} = -ie T_\rho \left[\gamma_n \partial/\partial J \right] (I_{12 \ldots \overline{n-1}}) G_{12 \ldots \overline{n-1}} - ie T_\rho \left[\gamma_n G_n \partial/\partial J \right] (I_{12 \ldots n} G_{12 \ldots n}),$$

and $\hspace{11cm} (3 \cdot 10)$

$$I_{12 \ldots n} = -ie T_\rho \left[\gamma_n \partial/\partial J \right] I_{12 \ldots \overline{n-1}} - ie T_\rho \left[\gamma_n G_n \partial/\partial J \right] I_{12 \ldots n}$$

$$+ ie T_\rho \left[\gamma_n G_n I_{12 \ldots n} \sum_{m=1}^{n} G_m \, \partial/\partial J \right] G_m^{-1} - ic T_\rho \left[\gamma_n G_n I_{12 \ldots n} \frac{G_{12 \ldots n}}{n!} \sum_{m=2}^{n} \sum_{(i_1 \ldots i_m)}^{(n)} \right.$$

$$\left. \cdot G_{i_{m+1} \ldots i_n}^{-1} \partial/\partial J \right] I_{i_1 \ldots i_m} + ic T_\rho \left[\gamma_n G_n I_{12 \ldots n} \frac{G_{12 \ldots n}}{n!} \sum_{(i_1 \ldots i_{n-1})}^{(n)} \right.$$

$$\left. \cdot \left\{ \sum_{m=1}^{n-2} \sum_{(i_1 \ldots i_m)}^{(n-1)} G_{i_1 \ldots i_{n-1}}^{-1} I_{i_1 \ldots i_m, i_n} + I_{12 \ldots n} \right\} G_{i_n} \, \partial/\partial J \right] G_{i_n}^{-1}$$

$$\underbrace{\hspace{4cm}}_{(n-m-1) \text{ factors except } (i_1 \ldots i_m)} \hspace{3cm} (3 \cdot 11)$$

in the same way as we followed on deriving the equations $(2 \cdot 14)$ and $(2 \cdot 17)$.

§4. Removal of the divergences in the quantities of two-body problem

We have thus far continued our analysis on the side of the S-matrix theory which is expanded into the power series of the coupling constant. So it is founded upon more restricted basis than Schwinger's[5], and therefore unfavourable pertaining to this point, but still useful in the removal of divergences.

Gupta-Dyson's[13] formalism is here adopted.* Let us distinguish unrenormalized quantities from renormalized ones by attaching star-symbol to them. Then we write the Lagrangian density as

$$L = - (1/4) F_{\mu\nu} F_{\mu\nu} - \psi(-i\gamma\partial + m)\psi - (1/2)(\partial A_\mu/\partial x_\mu)^2$$

$$+ (J_\mu + e\bar{\psi}\gamma_\mu\psi) A_\mu + \delta m \bar{\psi}\psi + (f/4) F_{\mu\nu} F_{\mu\nu}, \hspace{2cm} (4 \cdot 1)$$

where

$$e^* A_\mu^* = e A_\mu, \quad A_\mu^* = \theta A_\mu, \quad J_\mu^* = \theta J_\mu, \quad \theta = e^*/e, \quad f = 1 - \theta^{-2}. \hspace{1cm} (4 \cdot 2)$$

The Hamiltonian in the interaction representation corresponding to $(4 \cdot 1)$ is

$$H = -(f/4) F_{kl} F_{kl} - (f'/2) F_{k4}^2 - \delta m \bar{\psi}\psi - A_\mu(J_\mu + e\bar{\psi}\gamma_\mu\psi), \quad f' = f/(1-f), \hspace{0.5cm} (4 \cdot 3)$$

where for the Greek subscripts it is assumed to run from 1 to 4, whereas for the Latin from 1 to 3.

As regards the quantities in one-body problem, the modifications due to the additive terms in the Hamiltonian $(4 \cdot 3)$ have already been discussed by Utiyama et al.[11] Their fundamental equations are

---

* G. Takeda[5] has reformulated the renormalization theory by using the transformations of Lagrangian.

$$\{\gamma(p-e\langle A\rangle)+m+\textstyle\sum^*-\delta m\}\,G=1, \qquad (\square\,\delta_{\mu\rho}-L_{\mu\rho})\mathfrak{T}_{\rho\nu}=-1,$$

$$\{\square\,\delta_{\mu\nu}+f(\partial^2/\partial x_\mu\partial x_\nu-\delta_{\mu\nu}\square)\}\langle A_\nu(x)\rangle=-\{J_\mu(x)+iet_r(\gamma^\mu G(x,x))\},$$

$$P_{\mu\rho}=-ie^2t_r(\gamma^\mu G\Gamma^\rho G), \qquad L_{\mu\rho}=P_{\mu\rho}+f(\delta_{\mu\rho}\square-\partial^2/\partial x_\mu\partial x_\rho). \tag{4.4}$$

On the other hand, the modified equation of two-body problem is written as

$$\{\,[\gamma(p-e\langle A\rangle)+m+\textstyle\sum^*-\delta m]_1\,[\gamma(p-e\langle A\rangle)+m+\textstyle\sum^*-\delta m]_2-I_{12}\}\,G_{12}=1_{12} \tag{4.5}$$

and this is easily assured if one performs the same analysis as that of § 2 with modified Hamiltonian (4.3). Various quantities described above still include the divergences though the mass renormalization and a part of the charge renormalization have been carried out. In order to remove these remaining divergences, the following transformation is performed[*];

$$G=aG', \quad e=e', \quad \langle A\rangle=\langle A\rangle', \quad J=J, \quad \mathfrak{T}=\mathfrak{T}',$$

$$P=aP', \quad \Gamma=\Gamma'/a, \quad \textstyle\sum^*=\sum^{*\prime}, \quad \varLambda=\varLambda', \tag{4.6}$$

where a is an undetermined constant which is a function of e'. And Utiyama et al. have actually demonstrated the possibility of determining the constant a in each power of the coupling constant e' so as to make G', $\mathfrak{T}'$, I'' and $\langle A\rangle'$ divergence-free. Accordingly, the main task which will be mentioned below is to prove that those quantities peculiar to two-body problem are really divergence-free with thus chosen a.

Now, divide the quantities of two-body problem into those parts, any one of which is distinguished from another by the photon number exchanged between two open polygons of electron lines.

$$G_{12}=\sum_{n=0}^{\infty}G_{12}^{(n)}, \qquad I_{12}=\sum_{n=1}^{\infty}I_{12}^{(n)}. \tag{4.7}$$

And (4.5) turns into

$$\{G_1^{-1}G_2^{-1}-\sum_{n=1}^{\infty}I_{12}^{(n)}\}\sum_{m=1}^{\infty}G_{12}^{(m)}=1_{12}. \tag{4.5}'$$

When this equation is divided in conformity with the separation (4.7), we find

$$G_1^{-1}G_2^{-1}G_{12}^{(0)}=1_{12},\ \cdots,\ \ G_1^{-1}G_2^{-1}G_{12}^{(n)}-\sum_{m=1}^{n}I_{12}^{(m)}G_{12}^{(n-m)}=0. \tag{4.8}$$

Considering the transformations of I_{12} and G_{12}, namely

---

[*] This transformation is identical with the one given by Dyson[6]-Salam[14] in their treatments of S-matrix, but different from the one given by Utiyama et al.[11] The difference between them arises from the treatment of the so-called 'b-divergences'. We compare them with each other in the appendix 1.

[**] Consider the division-line which has the following properties in any diagram.

1) It crosses neither of two electron-open-polygons.
2) It crosses the least numbers of photon-lines exchanged between two electron-open-polygons.
3) If there are several ways to draw the division-line under the above two conditions, we take the nearest one to the electron-open-polygon which leads to the vertex x_2.

Denoting the number of photon-lines which this division-line crosses by n, we can classify every diagram by n.

$$G_{12}^{(n)} = a^{r_n+4} G_{12}^{\prime(n)}, \qquad I_{12}^{(n)} = a^{r_n} I_{12}^{\prime(n)}, \tag{4·6}'$$

we fix the values of the undetermined constants, $a^{r_1} \cdots a^{r_n}$, so as to make a set of equation $(4·8)$ invariant under the transformation $(4·6)'$. Substituting $(4·6)'$ into $(4·8)$, the following relation,

$$G_{12}^{(n)} = a^{r_n+4} G_{12}^{\prime(n)} = a^2 G_1' G_2' \Big(\sum_{m=1}^{\blacksquare} a^{r_m+r_{n-m}+4} I_{12}^{\prime(m)} G_{12}^{\prime(n-m)} + a^{r_n+2} I_{12}^{\prime(n)} G_{12}^{\prime(0)} \Big)$$

holds, as is easily confirmed. Therefore,

$$r_n = r_m + r_{n-m} + 2, \qquad 0 < m < n. \tag{4·9}$$

Now putting $r_1 = r$, we obtain $r_n = r_{n-1} + r_1 + 2 = n(r+2) - 2$. Though these values of r's are fixed by using only those parts of $(4·9)$ which corresponds to $m = n-1$, they satisfy $(4·9)$ for $m \neq n-1$, too; for r_n includes only the linear function of n and the constant term. From these results, $(4·8)$ becomes

$$-a^{n(r+2)} \sum_{l=0}^{n} I_{12}^{\prime(n-l)} G_{12}^{\prime(l)} = 1_{12} \hat{\delta}_{n,0}, \tag{4·8}'$$

where $I_{12}^{\prime(0)} = -G_1^{\prime-1} G_2^{\prime-1}$. $(4·5)'$ turns into

$$-\sum_{n=0}^{\infty} a^{n(r+2)-2} I_{12}^{\prime(n)} \sum_{m=0}^{\infty} a^{m(r+2)+2} G_{12}^{\prime(m)} = 1_{12}$$

and further into

$$-\sum_{n=0}^{\infty} \sum_{m=0}^{\infty} I_{12}^{\prime(n)} G_{12}^{\prime(m)} - \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \big(a^{(n+m)(r+2)} - 1 \big) I_{12}^{\prime(n)} G_{12}^{\prime(m)} = 1_{12}.$$

The second term vanishes because of $(4·8)'$. Thus we are left with

$$\big[G_1^{\prime-1} G_2^{\prime-1} - I_{12}' \big] G_{12}' = 1_{12}$$

where

$$I_{12}' = \sum_{m=1}^{\infty} I_{12}^{\prime(m)}, \qquad G_{12}' = \sum_{m=0}^{\infty} G_{12}^{\prime(m)}. \tag{4·10}$$

When we solve this equation in the restricted case of the scattering, we get

$$G_{12}'(x_1 x_2, \ x_1' x_2') = \{ G'(x_1, \ x_1') G'(x_2, \ x_2') - G'(x_1, \ x_2') G'(x_2, \ x_1') \}$$

$$+ \sum_{m=1}^{\infty} \int (G_1' G_2' I_{12}')^m (x_1 x_2, y_1 y_2) \cdot \{ G'(y_1, \ x_1') G'(y_2, \ x_2') - G'(y_1, \ x_2') G'(y_2, \ x_1') \}$$

$$\cdot dy_1 dy_2. \tag{4·11}$$

Now we assume that the series of $(4·11)$ is convergent. Taking the n-th term in $(4·11)$ and writing it down in the momentum space, we get

$$(k_1' k_2' | (G_{12})_n | k_1^i k_2^i) = \int (k_1^f | G' | k_1^i) (k_2^f | G' | k_2^i) \cdot$$

$$I_{12}'(k_1' k_2', \ k_1'' k_2'') (k_1'' | G' | k_1''') (k_2''' | G' | k_2''')$$

$$\times I'_{12}(k_1'''k_2''',\; k_1^{(4)}k_2^{(4)})\cdots(k_1^{(2n-2)}|G'|k_1^{(2n-1)})\,(k_2^{(2n-2)}|G'|k_2^{(2n-1)})\cdot$$

$$I'_{12}(k_1^{(2n-1)}k_2^{(2n-1)},\; k_1^{(2n)}k_2^{(2n)})$$

$$\times\{\,(k_1^{(2n)}|G'|k_1^i)\,(k_2^{(2n)}|G'|k_2^i)-(k_1^{(2n)}|G'|k_2^i)\,(k_2^{(2n)}|G'|k_1^i)\,\}\cdot$$

$$d^4k_1'd^4k_1''\cdots d^4k_1^{(2n)}\; d^4k_2'd^4k_2''\cdots d^4k_2^{(2n)}. \qquad (4\cdot12)$$

It is already known that $G_1'(G_2')$ is divergence-free. For some time, assuming that $I'_{12}(k_1'k_2',\; k_1''k_2'')$ is divergence-free, we are left with only the question as to whether the divergences do or do not arise from the integrals joining G_1' and G_2' to I'_{12}. In order to investigate it, I'_{12} and G' are expanded into the powers of the external current

$$I'_{12}(k_1'k_2',\; k_1''k_2'')=\sum_{m=0}^{\infty}\int I'_{12,m}(k_1'k_2'p_1p_2\cdots p_m)J'(p_1)J'(p_2)\cdots J'(p_m)$$

$$\times\delta(k_1'+k_2'+\sum_{i=1}^{m}p_i-k_1''-k_2'')d^4p_1\cdots d^4p_m,$$

$$(k_1'|G'|k_1'')=\sum_{m=0}^{\infty}\int G'_{,m}(k_1'p_1p_2\cdots p_m)J'(p_1)J'(p_2)\cdots J'(p_m)$$

$$\delta(k_1'+\sum_{i=1}^{m}p_i-k_1'')d^4p_1\cdots d^4p_m.$$

Here $J'(p)$ is the Fourier-coefficient of $J'(x)$ and δ-factor results from the conservation of energy and momentum. Since both I'_{12} and G' are free from divergences, the contributions from the terms with $m\geq1$ can be made not to give rise to any divergence, provided that the external current is properly chosen. This means that the external current may be chosen in such a way that the amplitude $J'(p)$ decreases rapidly enough to make the integral under consideratson converge. On the other hand, the term with $m=0$ of I'_{12} and $G_1'(G_2')$ does not contain the factor $J'(p)$. However, in this case G' corresponds to Dyson's S_F' and $k_{1(2)}^{(2m)}=k_{1(2)}^{(2m-1)}$. If we make the variable transformation, $k_1''=k_1'''+t,\; k_2''=k_2'''-t$, in the integral

$$\int\cdots\delta(k_1'+k_2'-k_1''-k_2'')\,I'_{12,0}(k_1'k_2',\; k_1''k_2'')\,S_F(k_1'')\,S_F(k_2'')$$

$$\times I'_{12,0}(k_1''k_2'',\; k_1'''k_2''')\delta(k_1''+k_2''-k_1'''-k_2''')\cdots,$$

this integral turns into

$$\int\cdots\delta(\cdots)\,I'_{12,0}(k_1'k_2',\; k_1'''+t,\; k_2'''-t)S_F'(k_1'''+t)\,S_F'(k_2'''-t)$$

$$\times I'_{12,0}(k_1'''+t,\; k_2'''-t,\; k_1''',k_2''')d^4t\cdots.$$

Accordingly, considering that I'_{12} has at most the same t-dependence as $D_F'(t)$, it is confirmed according to Dyson's criterium that the integration over t does not give rise to any divergence. Thus G'_{12}, too, turns out to be divergence-free except its convergence as the series. So the remaining task is to show that I'_{12} is really free from the divergence.

Finiteness of I'_{12}

We divide $(2\cdot17)$ in accordance with the separation $(4\cdot7)$, then $(2\cdot17)$ can be written as*

$$\underset{a}{I_{12}'^{(1)}} = -ic'^2 T_{\rho}[\gamma_1 \gamma_1 I_{12}''\mathfrak{T}'] -ic'^2 T_{\rho}[\gamma_1 G_1'\bar{\partial}/c'\partial J'] \underset{a}{\frac{I_{12}'^{(1)}}{}} + ic'^2 T_{\rho}\Big[\gamma_1 G_1' \underset{a}{\frac{I_{12}'^{(1)}}{}} G_i'\partial/c'\partial J'\Big]G_1'^{-1},$$

$$\vdots$$

$$\underset{a}{I_{12}'^{(n)}} = -ic'^2 T_{\rho}[\gamma_1 G_1' \sum_{m=1}^{n-1} (1_{12}+I_{12}'G_{12}')^{(n-1-m)}/2\cdot\partial^*/c'\partial J'] I_{12}'^{(m)}/a$$

$$+ ic'^2 T_{\rho}\Big[\gamma_1 G_1' \sum_{m=1}^{n-1} \frac{I_{12}'^{(m)}}{a}(1_{12}+G_{12}'I_{12}')^{(n-1-m)}/2\cdot\sum_{i=1}^{2} G_i'\partial^*/c'\partial J'\Big]G_i'^{-1}$$

$$- ic'^2 T_{\rho}[\gamma_1 G_1' \sum_{m=1}^{n} (1_{12}+I_{12}'G_{12}')^{(n-m)}/2\cdot\bar{\partial}/c'\partial J'] I_{12}'^{(m)}/a$$

$$+ ic'^2 T_{\rho}\Big[\gamma_1 G_1' \sum_{m=1}^{n} \frac{I_{12}'^{(m)}}{a}(1_{12}+G_{12}'I_{12}')^{(n-m)}/2\cdot\sum_{i=1}^{2} G_i'\bar{\partial}/c'\partial J'\Big]G_i'^{-1}, \qquad (4\cdot13)$$

when we take $(4\cdot6)'$ into account and choose $r=-2$.

By means of the mathematical induction, it is shown that I_{12}' is free from divergences through the following two steps; in $1)''$, it is shown that $I_{12}'^{(1)}$ is free from divergences, and in $2)^0$, it is verified that $I_{12}'^{(n)}$ is also divergence-free when $I_{12}'^{(1)}\cdots I_{12}'^{(n-1)}$ are assumed to be finite. $1)^0$ We can first show that $I_{12}'^{(1)} = -ic'^2 T_{\rho}[\Gamma_1'\Gamma_2'\mathfrak{T}']$ is the solution of the first equation of $(4\cdot13)$. That is, the second term of the right hand is transformed into

$$-ic'^2 \int \gamma^{\rho}(\xi_1) G_1(-ic'^2)\partial A_{\mu}'(\xi_1)/c'\partial J_{\nu}'(\xi_2)\cdot\mathfrak{T}_{\mu\rho}'(\xi_1'-\xi_1) \Gamma_{\nu}'(\xi_2)d^4\xi_1'd^4\xi_2 d^4\xi_1$$

$$-ic'^2 \int \gamma^{\rho}(\xi_1) G_1(-ic'^2)\Gamma_{\mu}'(\xi_1)\partial\mathfrak{T}_{\mu\rho}'(\xi_1'-\xi_1)/c'\partial J_{\nu}'(\xi_2)\cdot\Gamma_{\nu}'(\xi_2)d^4\xi_1'd^4\xi_2 d^4\xi_1,$$

where we put $I_{12}'^{(1)} = -ic'^2 T_{\rho}[\Gamma_1'\Gamma_2'\mathfrak{T}']$, and use the relations,

$$\partial\mathfrak{T}_{\sigma\rho}'(\eta_1-y_1)/\partial J_{\mu}'(y_2) = \partial\mathfrak{T}_{\sigma\mu}'(\eta_1-y_2)/\partial J_{\rho}'(y_1),$$

$$\int \mathfrak{T}_{\sigma\rho}'(\eta_1-y_1)\partial A_{\sigma}'(a_1 y_1', \eta_1)/\partial J_{\mu}'(y_2)\cdot d^4\eta_1$$

$$= \int \mathfrak{T}_{\sigma\mu}'(\eta_1'-y_2)\partial A_{\sigma}'(a_1 y_1', \eta_1)/\partial J_{\rho}'(y_1)\cdot d^4\eta_1.$$

And the third term is written as

$$-ic'^2 \int \gamma^{\rho}(\xi_1)\partial G_1'/c'\partial J_{\nu}'(\xi_2)\cdot(-ic'^2) \Gamma_{\mu}'(\xi_1)\mathfrak{T}_{\mu\rho}'(\xi_1'-\xi_1)d^4\xi_1'.$$

Hence considering $\sum^{*'} = ic'^2 T_{\rho}[\gamma G'\Gamma'\mathfrak{T}']$, we obtain

---

* $\partial/\partial J' = \partial^*/\partial J' + \bar{\partial}/\partial J'$, where $\partial^*/\partial J'$ stands for the differentiation which results in the increase of one exchanged photon between two electrons, while $\bar{\partial}/\partial J'$ stands for the differentiation which contributes to the radiative correction of the same electron.

$$R.H.S.=-ie'^{2}T_{p}[\gamma_{1}\Gamma_{2}\mathfrak{T}']-ie'^{2}\int \delta \sum_{1}^{\prime *}/(-)e'\partial J_{v}'(\xi_{2})\cdot\Gamma_{2}'^{v}(\xi_{2})d^{4}\xi,$$

$$=-ie'^{2}T_{p}[\Gamma_{1}'\Gamma_{2}'\mathfrak{T}'].$$

Since I'', $\mathfrak{T}'$ are finite and there arise no divergences from the integration-operation T_{p}, $I_{12}'^{(1)}$ is finite.

2)⁰ We introduce the notation $\bar{\partial}/\partial J'=\partial'/\partial J'+\partial''/\partial J'$, where $\partial'/\partial J'$ stands for the differentiation which gives rise to the term with the divergent constant of the vertex part, and $\partial''/\partial J'$ stands for the differentiation which gives rise to only finite terms. From the R.H.S. of (4·13), we divide the graphs of $I_{12}'^{(n)}$ into three groups. First group consists of those graphs which are built up from less than $I_{12}'^{(n)}$. The 1-st and 2-nd term, and the 3-rd and 4-th terms except for $m=n$ on the R.H.S. of (4·13) come under this group. Second group includes those graphs which are built up from simpler graphs of $I_{12}'^{(n)}$ and correspond to the 3-rd term with $m=n$, $\partial''/\partial J'$ and the 4-th term with $m=n$. The third group of diagrams is the 3-rd term with $m=n$ and $\partial'/\partial J'$, namely,

$$-ie'^{2}T_{p}[\gamma_{1}G_{1}'(1_{12}/2)\partial'/e'\partial J']I_{12}'^{(n)}/a.$$

Here all graphs of $I_{12}'^{(n)}$ are substituted for the expression $I_{12}'^{(n)}$ of this formula. The divergences included in this term can be settled in the following way. According to Salam's method, we divide the graphs of this group into $[2], [3], \cdots [n], \cdots$ corresponding to each step of iteration. $[1]$ is set aside for the diagrams of the first and second groups. Then we have the following set of equations.

$$[1]=I(1),$$
$$[2]=L_{1}\times[1]+I(2),$$
$$[3]=L_{1}\times[2]+L_{2}[1]+I(3),$$
$$\vdots$$
$$[n]=L_{1}\times[n-1]+L_{2}\times[n-2]+\cdots+L_{n-1}\times[1]+I(n).$$
$$\vdots \qquad (4\cdot14)$$

Here L_{1}, L_{2},$\cdots$ are the divergent constants of the vertex part. $I(2)$, $I(3)$,$\cdots$ have, respectively, the form that is made by substituting the finite part of the vertex part in each order for the first γ-factor in the expression $I(1)$. The summation of eqs. (4·14) becomes

$$[1]+[2]+[3]+\cdots+[n]=(1-L)^{-1}\cdot\{I(1)+I(2)+\cdots+I(n)+\cdots\},$$

$$L=L_{1}+L_{2}+\cdots+L_{n}+\cdots.$$

$[1]+[2]+[3]+\cdots+[n]+\cdots$ includes all the graphs of $I_{12}'^{(n)}$ and never includes the same graph more than once. We, therefore, obtain

$$I_{12}'^{(n)}/a=(1-L)^{-1}\cdot\{I(1)+I(2)+\cdots+I(n)+\cdots\}. \qquad (4\cdot15)$$

Then those parts of $I_{12}'^{(n)}$, which result from (4·15) by substituting the diagram of the first group for $I(1)$, prove to be finite. For $I_{12}'^{(1)}\cdots I_{12}'^{(n-1)}$ are finite by the

assumption and no divergence occurs in the integration involved. Next, those parts of $I_{12}^{\prime(n)}$, which result from $(4 \cdot 15)$ by substituting the diagrams now proved finite for $I(1)$, become also finite. In this way, $I_{12}^{\prime(n)}$ turns out to be finite up to the requisite order. So $I_{12}^{\prime}$ has no divergence in it provided that the convergence of the series, $I_{12}^{\prime} = \sum\limits_{n=1}^{\infty} I_{12}^{\prime(n)}$ is assumed.

§ 5.   Removal of the divergences in the quantities of n-body problem

We think of the removal of the divergences in those quantities peculiar to n-body problem.   The procedures followed in this section are quite similar to those in § 4.   So we shall mention only the points of the arguments.

G_N and I_N^* are divided into those parts, any one of which is distinguished from another in terms of the number** of the photon lines exchanged between a group of $(n-1)$ electrons and the n-th electron, namely,

$$G_N = \sum_{n=0}^{\infty} G_N^{(n)}, \qquad I_N = \sum_{n=1}^{\cdot} I_N^{(n)} .$$

Then $(3 \cdot 7)$ becomes

$$\left\{ G_{N-1}^{-1} G^{-1} - \Big(\sum_{M=1}^{n-2} \sum_{M} G_{N-M-1}^{-1} \sum_{n=1}^{\infty} I_{M+1}^{(n)} + \sum_{n=1}^{\infty} I_N^{(n)} \Big) \right\} \sum_{m=0}^{\infty} G_N^{(m)} = 1_N. \tag{5.1}$$

When we consider the transformation

$$G_N^{(n)} = a^N G_N^{(n)}, \qquad I_N^{(n)} = a^{-N} I_N^{(n)} \tag{5.2}$$

with $(4 \cdot 6)$, we obtain

$$G_{N-1}^{\prime-1} \, G^{\prime-1} \cdot G_N^{\prime(0)} = 1_N,$$
$$G_{N-1}^{\prime-1} \, G^{\prime-1} \, G_N^{\prime(n)} - \Big(\sum \sum G_{N-M-1}^{\prime-1} \sum_{m=1}^{n} I_{M+1}^{\prime(m)} + \sum_{m=1}^{n} I_N^{\prime(m)} \Big) G_N^{\prime(n-m)} = 0. \tag{5.3}$$

Accordingly, from $(5 \cdot 1)$ we find

$$G_N^{\prime} = \sum_{n=0}^{\infty} G_N^{\prime(n)}, \qquad I_N^{\prime} = \sum_{n=1}^{\infty} I_N^{\prime(n)},$$

$$\left\{ G_{N-1}^{\prime-1} G^{\prime-1} - \Big(\sum_{M=1}^{n-2} \sum_{M}^{(n-1)} G_{N-M-1}^{\prime-1} I_{M+1}^{\prime} + I_N^{\prime} \Big) \right\} G_N^{\prime} = 1_N. \tag{5.4}$$

As $(5 \cdot 4)$ is solved in the restricted case of the scattering, it follows that

$$G_N^{\prime} = \sum_{m=0}^{\infty} \left\{ G_{N-1}^{\prime} \, G^{\prime} \Big(\sum_{M=1}^{n-2} \sum_{M}^{(n-1)} G_{N-M-1}^{\prime-1} I_{M+1}^{\prime} + I_N^{\prime} \Big) \right\}^m G_{N-1}^{\prime} G^{\prime} \cdot 1_N. \tag{5.5}$$

Here under the assumption that various quantities peculiar to $(n-m)$-body problem

---

* For short, we write the n-electron Green-function as G_N and the n-electron interaction term as I_N.
** This photon-number is defined in a similar way to the one at the foot-note of p. 169.

$(m < n)$ are divergence-free, we investigate the n-body Green-function G'_N is really divergence-free or not. The quantities on the R.H.S. of $(5 \cdot 5)$ are assumed to be finite except I'_N. Further we assume that the series of $(5 \cdot 5)$ is convergent. Now when we assume for some time that I'_N is divergence-free, it is the integration connecting the factors, I'_N, $G'_{N-1}, \cdots$ etc., with one another that is likely to diverge. To investigate it, we had better expand the quantities, I'_N, $G'_{N-1}, \cdots$ etc., into the powers of the external current J'_μ and transform them into the momentum space. On that occasion, in case the external current $J'(p)$ is involved in the integrand, the integral can be made convergent provided that the external current $J'(p)$ is properly chosen.

On the other hand, in case the external current is not involved in the integrand, $\overset{\circ}{G}'(k)$ corresponds to Dyson's $S'_F(k)$, and the dependence on $k_1 \cdots k_N$ of $\overset{\circ}{G}'_{N-1}(k_1 \cdots k_{N-1})$ and $\overset{\circ}{I}'_N(k_1 \cdots k_N)$ cannot exceed that of $S'_F(k_1) \cdots S'_F(k_{N-1})$ and of $D'_F(k_1) \cdots D'_F(k_N)$, respectively. Therefore, any divergence does not appear in this integration. Hence, if we can only show that I'_N is free from divergences, G'_N turns out to be finite apart from the convergence of the series $(5 \cdot 5)$. As for the finiteness of $I'^{(l)}_N$, we can show it up to the requisite order of e' by the arguments similar to those made in $(4 \cdot 14)$. Hence, I'_N becomes finite if the convergence of the series, $I'_N = \sum_{l=1}^{\infty} I'^{(l)}_N$, is assumed.

§ 6. Conclusion

Though we assume the expandability in powers of the coupling constant throughout the above analysis, the equations in § 2 and § 3 can be derived independently of this assumption as is already shown by Schwinger.[5] As for the removal of the divergences, we have, however, no prescription to deal with it saving that by Dyson which is resorted to the perturbation theory. And if we follow this, we are inevitably to face the problem whether the power series with respect to the coupling constant does converge or not after the renormalization. This is quite a difficult problem with which much yet remains to be done in succession to Dyson's pioneer-work. However, for some time assuming that the series converges, all the quantities in § 2 and § 4 are easily understood in terms of the familiar quantities in one-body problem, and turn out to be divergence-free. But specially when we deal with the bound states, we can no more utilize the series expansion. Consequently, it is desirable to find a well-behaved solution of the system of integro-differential equations $(2 \cdot 4)$, $(2 \cdot 14)$, $(2 \cdot 16)$ etc. without resort to the series expansion.

Acknowledgement

Appendix 1.

As is well known in the S-matrix theory, there arises the appearance of the extra factor $z_1^{-1}(=a^{-1}$ in our notation) in the calculation of $\sum{}^{*}$ or $\mathit{\Gamma}_{\mu\rho}$, which is due to the divergences arising from the overlapping integral. (These divergences are called 'b-divergences'.) Salam[1] has shown that it is vital that this factor should appear in order to demonstrate the possibility of the renormalization. We consider a-transformation, following his prescription,

$$G=aG', \qquad e=a^{n}e', \qquad \langle A\rangle=a^{m}\langle A\rangle', \qquad f=a^{m}f',$$

$$\sum{}^{*}=a^{l}\sum{}^{*'}, \qquad \mathit{\Gamma}=a^{k}\mathit{\Gamma}'', \qquad A=a^{j}A', \qquad P=a^{h}P'.$$

$$(f=f') \tag{A$\cdot$1$\cdot$1}$$

$$\{D-e\gamma\langle A\rangle+\sum{}^{*}-\delta m\}\,G=1, \qquad (\because \ \gamma p+m=D),$$

$$\sum{}^{*}=ie^{2}T_{\rho}[\gamma G\mathit{\Gamma}\mathfrak{T}], \qquad \mathit{\Gamma}'=\gamma+A=-(1/e)\delta G^{-1}/\delta\langle A\rangle,$$

$$A=-(1/e)\delta\sum{}^{*}/\delta\langle A\rangle, \qquad \mathfrak{T}=\delta\langle A\rangle/\delta f,$$

$$P=-ie^{2}T_{\rho}[\gamma G\mathit{\Gamma}G], \qquad L_{\mu\rho}=P_{\mu\rho}+f(\delta_{\mu\rho}\square-\partial^{2}\,\partial x_{\mu}\partial x_{\rho}),$$

$$-(\square-L)\mathfrak{T}=1,$$

$$-\{\square\delta_{\mu\nu}+f(\partial^{2}/\partial x_{\mu}\partial x_{\nu}-\square\delta_{\mu\nu})\}\,\langle A_{\nu}\rangle=J_{\mu}+ie T_{\rho}[\gamma_{\mu}G]. \tag{A$\cdot$1$\cdot$2}$$

And the yet undetermined constants, n, m, l, k, j and h are fixed so that the relations 1)–5) may be kept invariant.

1) $\quad \mathit{\Gamma}=a^{k}\mathit{\Gamma}''=-a^{(1+n+m)}\delta G'^{-1}/e'\delta\langle A\rangle' \longrightarrow \mathit{\Gamma}'=\delta G'^{-1}/e'\delta\langle A\rangle' \quad (k=n+m+1),$

2) $\quad \sum{}^{*}=a^{l}\sum{}^{*'}=a^{2n+1+k}ie'^{2}T_{\rho}[\gamma G'\mathit{\Gamma}'\mathfrak{T}'] \longrightarrow \sum{}^{*'}=ie'^{2}T_{\rho}[\gamma G'\mathit{\Gamma}'\mathfrak{T}']$

$$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (l=2n+1+k=n-m),$$

3) $\quad P=a^{h}P'=-ie'^{2}a^{2n+2+k}T_{\rho}[\gamma G'\mathit{\Gamma}''G'] \longrightarrow P'=-ie'^{2}T_{\rho}[\gamma G'\mathit{\Gamma}''G']$

$$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (h=2n+2+k=n-m+1),$$

4) $\quad A=a^{j}A'=-a^{l-m-n}\delta\sum{}^{*'}/e'\delta\langle A\rangle' \longrightarrow A'=-\delta\sum{}^{*'}/e'\delta\langle A\rangle'$

$$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (j=l-m-n=-2m),$$

5) $\quad \mathfrak{T}=\mathfrak{T}'=\delta\langle A\rangle'/\delta f',$

6) $\quad \mathit{\Gamma}'=\gamma+A \longrightarrow \mathit{\Gamma}''=\gamma+(a^{-k}-1)\gamma+a^{n-m+1}A'$

7) $\quad -(\square\delta_{\mu\rho}-L_{\mu\rho})\mathfrak{T}_{\rho\nu}=1 \longrightarrow -(\square\delta_{\mu\rho}-L'_{\mu\rho})\mathfrak{T}'_{\rho\nu}=1,$

$$\qquad\qquad\qquad\qquad L'_{\mu\rho}=a^{h}P'_{\mu\rho}+f'(\delta_{\mu\rho}\square-\partial^{2}/\partial x_{\mu}\partial x_{\rho}).$$

We take here $h=1$, whereas Utiyama et al. $h=0$. (Their paper will be referred to

hereafter as (U).) Using the above relations, 1)–4), we get $n=m$, $k=-(2n+1)$, $l=0$ and $j=-2n$.

8) Now as we take $n=0$, we find $h=1$, $k=-1$ and $n=m=l=j=0$. We expand $\sum^{*\prime}$ in the powers of $\langle A \rangle'$. Then, we get

$$\sum\nolimits^{*\prime} = \{ \overset{\circ}{\sum}{}^{*\prime} - e' \overset{\circ}{A'} \langle A \rangle' + \sum\nolimits_F^{*\prime} \} / a \tag{A·1·3}$$

where

$$\overset{\circ}{\sum}\nolimits_n^{*\prime}(k) = A_n' + B_n(\gamma k + m) + \overset{\circ}{\sum}\nolimits_{f,n}^{*\prime}(k) \, ,,$$

and

$$\overset{\circ}{A}{}_n'(k,\ k') = L_n' \gamma \partial(k-k') + \overset{\circ}{A}{}_{f,n}'(k,\ k').$$

$\overset{\circ}{\sum}{}_F^{*\prime}$, $\overset{\circ}{\sum}{}_f^{*\prime}$ and $\overset{\circ}{A}{}_f'$ are all finite quantities. Accordingly,

$$\{ D - e\gamma \langle A \rangle + \sum\nolimits^* - \partial m \} G = 1 \longrightarrow DG' + K'G' = 1,$$

$$(k | K_n' | k') = \partial(k-k') [(A' - a\partial m)_n + (a-1+B)_n(\gamma k + m)]$$

$$- e' \sum_{j=0}^{n-1} (a-1+L)_{n-j-1} \gamma^\mu (k | \langle A_\mu \rangle_j' | k') + \partial(k-k') \overset{\circ}{\sum}\nolimits_{f,n}^{*\prime}(k) + \sum\nolimits_{F,n}^{*\prime}(k,\ k')$$

$$- e' \sum_{j=0}^{n-1} \{ \gamma^\mu + \overset{\circ}{A}{}_f'^\mu(k,\ k') \}_{n-j-1} (k | \langle A_\mu \rangle_j' | k').$$

By the way, it is to be remarked that our $\overset{\circ}{\sum}{}_F^{*\prime}$, $\overset{\circ}{\sum}{}_f^{*\prime}$ and $\overset{\circ}{A}{}_f'$ are not equal to the corresponding quantities in (U). Suppose that the unknown parameter a and ∂m are fixed by

$$a = 1 - \sum_{n=2}^{\infty} B_n e'^n = s_2(e'), \qquad a\partial m = \sum_{n=2}^{\prime} A_n' e'^n.$$

Then K' turns out to be finite. Further, we obtain from 4), 6) and (A·1·3),

$$\Gamma_{\mu,n}'(k,\ k') = \{ \gamma^\mu \partial(k-k') + (a-1+L)\gamma^\mu \partial(k-k') + \overset{\circ}{A}{}_{f,\mu}'(k,\ k') + A_F'^\mu(k,\ k') \}_{,,}$$

and this is finite because $B_n = L_n$. Lastly, we expand Γ'' in the powers of external current J'. It follows that

$$(k | P_n' | k') = (1/a) \{ \partial(k-k') \overset{\circ}{P}{}_n'(k) + (k | P_{n,F}' | k') \},$$

$$\overset{\circ}{P}{}_n'(k) = e'^n A_n(\partial_{\mu\nu} k^2 - k_\mu k_\nu). \tag{A·1·4}$$

Here $(k | P_{n,F}' | k')$ $P_{f,n}'$ are again finite. The extra factor $1/a$ just cancels the factor a which stands in front of P' in 7). Thus, supposing that f' is given by

$$f'(e') = \sum_{n=2}^{\infty} e'^n A_n = s_3(e'),$$

we get

$$(k | L_{\mu\nu}' | k')_n = (\overset{\circ}{P}{}_{\mu\nu}'(k))_{n,f} \partial(k-k') + (k | P_{\mu\nu,F}' | k')_n,$$

which proves to be finite. Accordingly, if we follow the same argument done in (U) concerning the further analyses, we can demonstrate the finiteness of G', $\mathfrak{T}'$ and $\langle A \rangle'$.

In short, the main difference between (U) and our argument lies in the appearance of the extra factor $1/a$ both in $(A\cdot1\cdot3)$ and $(A\cdot1\cdot4)$. In order to remove this factor which is due to the ' b-divergence ', we must take $h=1$ and $n=0$ in 7) and 8). As in (U), however, this factor does not appear, so $h=0$ and $n=-1/2$ (therefore $m=-1/2$, $k=l=j=-1$) are taken. According to (U), when the systems of the integro-differential equations are solved successively by iteration, the way in which $\sum^{*\prime}(P')$ may be built by the insertion of the vertex part is uniquely determined by 2) (or 3)). The order of integration is, therefore, fixed by the order of iteration in the integrals that involve divergences. Provided that the integration variables are made to tend to infinity by this order, there do not arise b-divergences. So the above stated extra factor $1/a$, too, does not appear. That is, the self-energy parts ($\sum^{*\prime}$ and P'), which are obtained by iteration in the form of the series, are summed up by the different ways in the two treatments ; in (U), both end-points are unsymmetrically dealt with, whereas in Salam's and ours, they are symmetrically dealt with. And their results ought to coincide with each other, provided that the series might be convergent. Unfortunately, the series is divergent and the two results are different by a divergent factor a.

References

1) G. Breit, Phys. Rev. **34** (1929), 553.
2) S. M. Dancoff, Phys. Rev. **78** (1950), 382.
   I. Tamm, J. Phys. (U.S.S.R) **9** (1945), 449.
3) E. E. Salpeter and H. A. Bethe, Phys. Rev. **84** (1951), 1232.
4) M. Gell-Mann and F. Low, Phys. Rev. **84** (1951), 350.
5) J. Schwinger, Proc. N. A. S. **37** (1951), 452.
6) F. J. Dyson, Phys. Rev. **75** (1949), 1736.
7) M. M. Lévy, Phys. Rev. **84** (1951), 441 ; **88** (1952), 725.
8) F. J. Dyson, Phys. Rev. **85** (1952), 631.
9) Y. Katayama and K. Yamazaki, Prog. Theor. Phys. **7** (1952), 601.
10) R. Utiyama and T. Imamura, Prog. Theor. Phys. **9** (1953), 431.
11) R. Utiyama et al., Prog. Theor. Phys. **8** (1952), 77.
12) G. C. Wick, Phys. Rev. **80** (1950), 268.
13) S. N. Gupta, Proc. Phys. Soc. (London) **A64** (1951), 426.
    F. J. Dyson, Phys. Rev. **83** (1951), 608.
14) A. Salam, Phys. Rev. **82** (1951), 217.
15) G. Takeda, Prog. Theor. Phys. **7** (1952), 359.

Bound States and S-matrix in Quantum Field Theory

Shinzo NAKAI

Department of Physics, Osaka University

(Received December 7, 1953)

The relationship between bound states and steady states defined by Dyson is investigated following Hamilton's treatment. We, however, obtain the different expression of S-matrix from that of Hamilton, when the relevant terms of the S-matrix for the two-body scattering are resolved into the infinite series. So from our own standpoint, a homogeneous integral equation describing bound states (i. e. Nambu and S. B.'s equation) is derived keeping close connection with Møller's S-matrix theory.

§ 1. Introduction

As is well known, the S-matrix[1] evaluates transitions between initial and final states which describe non-interacting particles. Here, the free particle state is supposed to describe the state of the system before the interaction is adiabatically switched on or after it is adiabatically switched off, respectively. The steady states of a system of particles are defined, according to Dyson[2], by

$$S(\infty)\,\Phi = \lambda\Phi\,,$$

where $|\lambda|^2 = 1$.

According to Møller[3], the energy values in bound states are to be derivable from the S-matrix, provided that the S-matrix is assumed to be an analytic function of the total energy W of the system. They are simply given by the zero points of the eigen-value λ of S on the real axis in the complex W_0-plane, the eigen-value W_0 of W being regarded as a complex variable. By the process of analytic continuation, the wave function ψ_{W_0} may then be given an unambiguous meaning also for complex values of the variable W_0 as well as for real values W_0 smaller than the minimum value $W^0{}_m$ of the energy in a continuous state. And ψ_{W_0} will be a solution of the Schroedinger equation and an eigen-function corresponding to a bound state only for those real values of $W_0 < W^0{}_m$ which make S equal to zero. In this case, the asymptotic expression of ψ_{W_0} for large values of the relative distance between the particles vanishes exponentially with increasing distance.

Kita[4] has already derived a homogeneous integral equation for bound states, keeping the close connection with the S-matrix theory along the above-stated Møller's prescription. He, however, started from Feynman[5] theory and thereby the connection between the wave function and the state function is less obvious without recourse to Gell-Mann and Low's work[6]. Further, the derivation of a homogeneous equation from an inhomogeneous equa-

tion for the scattering is quite artificial in his treatment. For the eigen-value λ of S is not involved in the inhomogeneous equation, untill this equation is divided by λ. And thereafter the pole of λ, which corresponds to the zero point in Møller's treatment, is investigated. The limiting $\lambda \to \infty$ makes the inhomogeneous term vanish and also makes the solution have the asymptotic expression which vanishes exponentially with the increasing mutual distance.

Further we know that an equation for the bound states of particles which interact through a quantized field has been derived by Salpeter and Bethe[7], and the fundamental basis of this equation has been investigated by Gell-Mann[8] and Low in terms of the method of Feynman. And the interesting problem to point out the simple relationship between S. B.'s result and the S-matrix has been developed by Hamilton[9]. But when the relevant terms of the S-matrix for the two-body scattering are resolved into the infinite series, it seems to the author that we should be led to the different expression from the one which Hamilton has derived. In fact, as the result of his calculation, the definition of bound states has no connection at all with the eigen-value λ of S, and this fact appears to run counter to Møller's prescription summarized above. It is the aim of this paper to follow Hamilton's treatment which starts from the steady state and to show that a homogeneous equation for bound states and an inhomogeneous equation for the scattering can be derived in a more natural way. For the eigen-value λ of S is originally included in the inhomogeneous equation. As the steady states are symmetric with respect to past and future, so the whole analysis can be carried out relativistically.

§ 2. Steady states and an inhomogeneous integral equation

The steady states of a system of particles are defined in terms of

$$S(\infty)\,\Phi = \lambda\Phi .\tag{2·1}$$

Now, $S(\infty, -\infty)$ can be written as

$$S(\infty, -\infty) = \sum_{n=0}^{\infty} (-i)^n/n! \int_{-\infty}^{\infty} d^4x_1 \cdots \int_{-\infty}^{\infty} d^4x_n \, T(H(x_1)\cdots H(x_n)),$$

where

$$H(x) = if\,\bar{\psi}(x)\,O_\lambda \psi(x)\,\phi_\lambda(x) + H_{\text{self}}.$$

And according to Wick's[5] theorem, it can be expanded into

$$S = 1 + \sum_{i,j=0}^{\infty} \int N(\bar{\psi}_\alpha(p_1')\cdots\bar{\psi}_\beta(p_i'), \psi_\tau(p_1)\cdots\psi_\delta(p_i), \phi_\lambda(q_1)\cdots\phi_\nu(q_j))$$

$$\times \delta^4(\text{energy-momentum})\; U^{i,j,}_{\alpha\ldots\beta,\tau\ldots\delta,\lambda\ldots\nu}\,(p_1'\cdots p_i', p_1\cdots p_i, q_1\cdots q_j)$$

$$\times d^4p_1'\cdots d^4p_i' d^4p_1\cdots d^4p_i d^4q_1\cdots d^4q_j ,\tag{2·2}$$

where the summation is carried out over all the possible numbers of i, j, which are allowed by the given H. The contributions from the isolated closed loops are neglected, because

they are dealt with by normalizing the probability that a vacuum remain vacuum as one. On the other hand, an arbitrary free-particle state may be expanded as

$$\Phi = \sum_{i,j,l} \int N(\bar{\psi}_\alpha(p') \cdots \bar{\psi}_\beta(p_i'), \psi_\tau(p_1) \cdots \psi_\delta(p_j), \phi_\lambda(q_1) \cdots \phi_\nu(q_l)) \, \Phi_{\text{vac}} .$$

$$\times \delta^4 (\text{energy-momenta}) \;\; \varphi^{ijl}_{\alpha \cdots \beta, \tau \cdots \delta, \lambda \cdots \nu} (p_1' \cdots p_i', p_1 \cdots p_j, q_1 \cdots q_l)$$

$$\times d^4 p_1' \cdots d^4 p_i' d^4 p_1 \cdots d^4 p_j d^4 q_1 \cdots d^4 q_l. \tag{2.3}$$

The wave functions φ^{ijl}'s for the steady state are given by the solution of (2.1), when (2.3) and (2.2) are substituted into (2.1). Total energy and momenta of the system of particles are constants of collision and therefore commute with S-matrix. So we can take the complete set of collision constants including the total energy-momenta for the specification of any steady state. From this fact, steady states can be constructed in (2.3) out of only those states which have definite total energy and momenta $(\boldsymbol{K}, W)$.

Now we restrict ourselves to two-body problem, and to the steady states with the definite total energy-momenta $(\boldsymbol{K}, W)$, $2M \le W < 2M + \mu$. We write

$$X_\mu = (x_1{}^\mu + x_2{}^\mu)/2, \quad x_\mu = x_2{}^\mu - x_1{}^\mu,$$

$$K_\mu = (k_1{}^\mu + k_2{}^\mu), \quad k_\mu = (k_2{}^\nu - k_1{}^\nu)/2,$$

$$K = (\boldsymbol{K}, W), \quad k = (\boldsymbol{k}, \epsilon). \qquad \gamma_\mu \gamma_\nu + \gamma_\nu \gamma_\mu = 2\delta.$$

The equation (2.1) takes the form*

$$\varphi(\boldsymbol{k}_1, \boldsymbol{k}_2) + \int (\boldsymbol{k}_1, \boldsymbol{k}_2 \mid U \mid \boldsymbol{k}_1' \boldsymbol{k}_2') d^3 k_1' d^3 k_2' \delta^3 (\boldsymbol{K}' - \boldsymbol{K}_0)$$

$$\times \delta(\sqrt{\boldsymbol{k'}_1{}^2 + M^2} + \sqrt{\boldsymbol{k'}_2{}^2 + M^2} \mp W_0) C \Lambda_1{}^\pm(\boldsymbol{k}_1') \Lambda_2{}^\pm(\boldsymbol{k}_2') \varphi(\boldsymbol{k}_1' \boldsymbol{k}_2') = \lambda \varphi(\boldsymbol{k}_1 \boldsymbol{k}_2),$$

where

$$\Lambda^+(\boldsymbol{k}_i) = (\gamma_4 \epsilon_i{}^* - i\gamma \boldsymbol{k}_i + M)/2M, \quad \Lambda^-(\boldsymbol{k}_i) = (\gamma_4 \epsilon_i{}^* - i\gamma \boldsymbol{k} - M)/2M,$$

$$C = -M^2/\epsilon_1{}^* \epsilon_2{}^*, \quad \epsilon_i{}^* = \sqrt{\boldsymbol{k}_i{}^2 + M^2} \quad (i = 1, 2).$$

The choice of $\pm$ signs attached to projection operators is made according as the system

---

* We analyze $\psi_\alpha(x)$ into Fourier components as

$$\psi_\alpha(x) = (2\pi)^{-3/2} \int d^3 k (M^2/\boldsymbol{k}^2 + M^2)^{1/4} \left[\sum_{r=1,2.} u_\alpha{}^r(\boldsymbol{k}) a_r(\boldsymbol{k}) e^{ikx} + \sum_{r=3,4.} u_\alpha{}^r(\boldsymbol{k}) b_r{}^*(\boldsymbol{k}) e^{-ikx} \right].$$

And the wave function $\varphi(\boldsymbol{k}_1 \boldsymbol{k}_2)$ of two ordinary particles is

$$\varphi(\boldsymbol{k}_1 \boldsymbol{k}_2) = [\varphi^{020}(\boldsymbol{k}_1 \boldsymbol{k}_2) - \varphi^{020}(\boldsymbol{k}_2 \boldsymbol{k}_1)]$$

and that of antiparticles is

$$\varphi(\boldsymbol{k}_1 \boldsymbol{k}_2) = [\varphi^{200}(\boldsymbol{k}_1 \boldsymbol{k}_2) - \varphi^{200}(\boldsymbol{k}_2 \boldsymbol{k}_1)] \qquad (U = U^{220}).$$

It is to be noted that the corresponding equation to (2.4) for the system of an ordinary particle and an antiparticle is the simultaneous one regarding the wave functions of two fermions and those of bosons only.

is consisting of two ordinary particles or two antiparticles, respectively. It is helpful to write all quantities four-dimensionally. Then, the above two equations can be united into the form

$$\varphi(k_1 k_2) + \int P'(k_1 k_2)(k_1 k_2 \mid U \mid k_1' k_2') d^4 k_1' d^4 k_2' \varphi(k_1' k_2') = \lambda \varphi(k_1 k_2), \qquad (2\cdot4)$$

in terms of

$$\varphi(k_1 k_2) = P'(k_1 k_2) \delta^4(k_1 + k_2 - K_0) \varphi(\mathbf{k}_1 \mathbf{k}_2)$$

and

$$P'(k_1 k_2) = (\gamma k_1 + iM)(\gamma k_2 + iM) \delta(k_1^2 + M^2) \delta(k_2^2 + M^2). \qquad (2\cdot5)$$

From now on, we shall simply write

$$S\varphi = (1 + P'U)\varphi = \lambda\varphi \qquad (2\cdot4')$$

instead of $(2\cdot4)$, considering that $(k_1 k_1 \mid P' \mid k_1' k_2') = \delta^4(k_1 - k_1') \delta^4(k_2 - k_2') P(k_1 k_2)$.

The graphs which describe the mutual scattering of two fermions can be classified according to the graphical method of Feynman and given by*

$$S = 1 + P'U,$$

$$U = I + IGU = I + I(1 + GI + GIGI + \cdots)GI,$$

$$I = \sum_{i=1}^{\infty} I_i. \qquad (2\cdot6)$$

Here, $I_1, I_2, \cdots I_n, \cdots$ stand for the proper-connected diagrams of two-body interaction and I_n, for example, includes all the proper-connected diagrams in which n bosons are exchanged between two fermions. By a proper-connected diagram in two-body problem, is meant such a diagram that cannot be divided into two parts by only two fermion lines. $G(k_1 k_2)$ is the product of Feynman propagation function, $S'_F(k_1).S'_F(k_2)$, relating to two fermions with energy-momenta k_1 and k_2 and with mass M. Let us now introduce the notation $P(k_1, k_2)$ by $G(k_1, k_2) = P(k_1, k_2) + P'(k_1, k_2)$, where $P'(k_1, k_2)$ is defined by $(2\cdot5)$.

---

* These relations are deduced by means of the method used in the analysis of two-body Green-function.[10] However, as the two-particle Green-function is defined by

$$G_{12}(x_1 x_1 x_1' x_2') = (\text{vac} \mid T(\psi(x_1)\psi(x_2)\bar\psi(x_1')\bar\psi(x_2').S) \text{ vac})/S \text{ vac},$$

we can analogize (2. 6) from the relations which G_{12} obeys. That is, the relations,

$$G_{12} = G' + G/G G_{12} = G' + G[IG' + IGIG_{12}],$$

$$\begin{cases} (k_1' k_2' \mid G' \mid k_1 k_2) = S_{F'}(k_1' - k_1) S_{F'}(k_2' - k_2) - S_{F'}(k_1' - k_2) S_{F'}(k_2' - k_1), \\ (k_1' k_2' \mid G \mid k_1 k_2) = S_{F'}(k_1' - k_1) S_{F'}(k_2' - k_2), \quad S_{F'}(k) = S_F(k)/g(k), \\ g(k) = 1 + \dfrac{k^2 + M^2}{M^2} f_1(k^2 + M^2) + \dfrac{i\gamma k + M}{2M} f_2(k^2 + M^2). \end{cases}$$

suggest that

$$S = 1 + P'U = 1 + P'(I + IGU).$$

Here, f_1 and f_2 are functions of $(k^2 + M^2)$.

The $P(k_1, k_2)$ factor will be written when both the fermion lines to which it refers do not describe free particle states. One can then divide U as

$$U = \sum_{n=0}^{\infty} U^{(n)}$$

where $U^{(n)}$ includes the P'-factor n times. $(2 \cdot 6)$ takes the following form:

$$U^{(0)} = I + IPU^{(0)}, \quad U^{(1)} = IP'U^{(0)} + IPU^{(1)}, \quad \cdots$$

$$\cdots, \quad U^{(n)} = IP'U^{(n-1)} + IPU^{(n)}, \quad \cdots \tag{2.7}$$

Putting $U^{(0)} = V$ and $U^{(n)} = VP'U^{(n-1)}$, eqs. $(2 \cdot 7)$ can be solved as

$$U^{(0)} = V = I + IPV = I + I(1 + PI + PIPI + \cdots)PI,$$

$$\cdot U^{(1)} = VP'V, \quad U^{(2)} = VP'VP'V, \cdots, U^{(n)} = \underbrace{VP'VP' \cdots VP'V}_{n \ P'\text{-factors}}.$$

It is, therefore, obvious that the relations[*]

$$S = 1 + P'V + P'VP'V + \cdots + \underbrace{P'V\ P'V \cdots P'V}_{n \ P'\text{-factors}} + \cdots$$

$$= 1 + P'VS,$$

$$S\varphi = \lambda\varphi = (1 + P'VS)\varphi = (1 + \lambda P'V)\varphi$$

hold. Hence,

$$P'V\varphi = (1 - \bar{\lambda})\varphi, \quad \bar{\lambda} = 1/\lambda. \tag{2.8}$$

In order to solve this equation $(2 \cdot 8)$, it is convenient to introduce an auxiliary function $\psi(k_1 k_2)$. It is defined by the relations

$$P'\psi = 0$$

and

$$\psi = (1 + PI + PIPI + \cdots)PI\varphi, \tag{2.9}$$

otherwise. Then, $(2 \cdot 8)$ and $(2 \cdot 9)$ become

$$P'I\varphi + P'I\psi = (1 - \bar{\lambda})\varphi$$

$$PI\varphi + PI\psi = \psi. \tag{2.10}$$

Adding these two equations and recalling the definition of the P-factor, we get

---

[*] The difference between Hamilton's result and ours appears in this expression of S-matrix. According to him, S can be written as

$$S = P'TP'TP'T \cdots = P'TS, \text{ where } P'T = 1 + P'V.$$

Hence,

$$S\varphi = \lambda\varphi = P'TS\varphi = \lambda P'T\varphi, \quad \therefore \varphi - P'T\varphi = (1 + P'V)\varphi, \quad \therefore P'V\varphi = 0.$$

However, it seems clear that the direct calculation leads to (2.8).

Our further arguments, therefore, become quite different from Hamilton's owing to this different point.

$$GI\theta = \theta - \bar{\lambda}\varphi \qquad (2\cdot11)$$

for the wave function

$$\theta = \varphi + \phi. \qquad (2\cdot12)$$

It can be said concerning the equations $(2\cdot10)$ that these two equations are complementary with each other; both sides of the first equation act to the left only on the wave function describing a free particle two-body state and those of the second act to the left on the wave function which describes both particles relating not to the free particle state. In fact, $\varphi(k_1k_2)$ or $\phi(k_1k_2)$ has the non-vanishing values only when both arguments refer to or do not refer to the free particle state, respectively. $\theta(k_1k_2)$ is thus defined in a whole region of k_1 and k_2. Operating on $(2\cdot11)$ by G^{-1} from the left gives

$$I\theta = G^{-1}\theta, \quad G^{-1} = (i\gamma k_1 + M)(i\gamma k_2 + M)g(k_1)g(k_2), \qquad (2\cdot13)$$

since there exists the relation $G^{-1}\varphi = 0$. This relation $G^{-1}\varphi = 0$, results from the fact that φ is the wave function describing the free particle state with the definite energy-momenta $(\boldsymbol{K}, W)$, $2M \leq W < 2M + \mu$. The wave function θ given by $(2\cdot12)$, therefore, obeys Nambu and S. B.'s equation.

We shall next derive an equation for the steady state with the energy-momenta $(\boldsymbol{K}, W)$, $2M + 2\mu > W \geq 2M + \mu$. This time, we need two terms, $\varphi(k_1k_2)$ and $\varphi(p_1p_2q)$, in $(2\cdot3)$ where $K = k_1 + k_2 = p_1 + p_2 + q$. k_1, k_2 and p_1, p_2 are energy-momenta of the two fermions and q those of the boson.

Taking four matrix elements,

$$(k_1k_2 \mid S_{22} \mid k_1'k_2'), \qquad (k_1k_2 \mid S_{23} \mid p_1' p_2'q'),$$

$$(p_1p_2q \mid S_{32} \mid k_1'k_2'), \qquad (p_1p_2q \mid S_{33} \mid p_1' p_2'q'),$$

the following relations are found in the same way as we used on deriving $(2\cdot8)$. Namely,

$$S_{22} = 1_{22} + P_2'V_{22}S_{22} + P_2'V_{23}S_{32},$$

$$S_{23} = \quad P_2'V_{22}S_{23} + P_2'V_{23}S_{33}, \qquad G_2 = P_2 + P_2',$$

$$S_{32} = \quad P_3'V_{32}S_{22} + P_3'V_{33}S_{32}, \qquad G_3 = P_3 + P_3',$$

$$S_{33} = 1_{33} + P_3'V_{32}S_{23} + P_3'V_{33}S_{33},$$

or simply

$$S = 1 + P'VS,$$

where

$$1 = \begin{pmatrix} 1_{22} & 0 \\ 0 & 1_{33} \end{pmatrix}, \quad P'V = \begin{pmatrix} P_2'V_{22} & P_2'V_{23} \\ P_3'V_{32} & P_3'V_{33} \end{pmatrix}, \quad S = \begin{pmatrix} S_{22} & S_{23} \\ S_{32} & S_{33} \end{pmatrix}. \qquad (2\cdot14)$$

And

$$V_{22}=I_{22}+I_{22}P_2V_{22}+I_{23}P_3V_{32},$$

$$V_{23}=I_{23}+I_{22}P_2V_{23}+I_{23}P_3V_{33},$$

$$V_{32}=I_{32}+I_{32}P_2V_{22}+I_{33}P_3V_{32},$$

$$V_{33}=I_{33}+I_{32}P_2V_{23}+I_{33}P_3V_{33}.$$

The equation $(2 \cdot 1)$ which determines the steady states becomes

$$P_2I_{22}\varphi_2+P_2I_{23}\varphi_3=(1-\bar{\lambda})\varphi_2,$$

$$P_3I_{32}\varphi_2+P_3I_{33}\varphi_3=(1-\bar{\lambda})\varphi_3, \quad (\bar{\lambda}=1/\lambda). \tag{2.15}$$

It is useful here, too, to introduce auxiliary functions defined by

$$P_2'\psi=0, \ P_3'\psi_3=0$$

and

$$\psi_2=P_2V_{22}\varphi_2+P_2V_{23}\varphi_3$$

$$\psi_3=P_3I_{32}\varphi_2+P_3V_{33}\varphi_3,$$

otherwise. $(2 \cdot 15)$ then turns into

$$\left. \begin{array}{cc} P_2I_{22}\theta_2+P_2I_{23}\theta_3=\psi_2, & P_3I_{32}\theta_2+P_3I_{33}\varphi_3=\psi_3 \\ P_2'I_{22}\theta_2+P_2'I_{23}\theta_3=(1-\bar{\lambda})\varphi_2, & P_3'I_{32}\theta_2+P_3'I_{33}\theta_3=(1-\bar{\lambda})\varphi_3 \end{array} \right\}, \tag{2.16}$$

where

$$\theta_2=\varphi_2+\psi_2, \ \theta_3=\varphi_3+\psi_3.$$

Making the sum of the first and the third equations, and likewise that of the second and the fourth, we obtain

$$\left. \begin{array}{c} G_2I_{22}\theta_2+G_2I_{23}\theta_3=\theta_2-\bar{\lambda}\varphi_2 \\ G_3I_{32}\theta_2+G_3I_{33}\theta_3=\theta_3-\bar{\lambda}\varphi_3 \end{array} \right\}. \tag{2.17}$$

Since $G_2^{-1}\varphi=0$ and $G_3^{-1}\varphi_3=0$, operating on $(2 \cdot 17)$ by G_2^{-1} in the first equation and G_3^{-1} in the second gives

$$\left. \begin{array}{c} G_2^{-1}\theta_2=I_{22}\theta_2+I_{23}\theta_3 \\ G_3^{-1}\theta_3=I_{32}\theta_2+I_{33}\theta_3 \end{array} \right\}. \tag{2.18}$$

This simultaneous equations will be regarded as the generalized type of equations of S. B. equation. In this manner, one can construct the similar equations in each case when the total energy W is given by $2M+2\mu\leq W < 2M+3\mu,\cdots, \ 2M+(n-1)\mu\leq W < 2M+n\mu,\cdots,$ respectively.

§ 3.   Bound states and a homogeneous integral equation

The integral equation which θ obeys has been derived as

$$G \cdot I\theta = \theta - \varphi/\lambda \,, \tag{3.1}$$

when the steady state function φ has the total energy W, $2M + \mu > W \geq 2M$.

We write now

$$\Psi(x_1 x_2) = \iint d^4k_1 d^4k_2 \exp \cdot \{i(k_1 x_1 + k_2 x_2)\} \cdot \theta(k_1 k_2). \tag{3.2}$$

Defining $\theta(k_1 k_2)$ by

$$\theta(k_1 k_2) = \delta^4(K - K_0) \theta(k_1 k_2),$$

one can write

$$\Psi(x_1 x_2)\Big|_{t_1 = t_2 = T/2} = e^{iK_0 X} \iint d\epsilon d^3k \{GI\theta + \bar{\lambda}\varphi\} e^{ikx} = e^{iK_0 X} \psi(x)\Big|_{t_1 = t_2}. \tag{3.3}$$

We assume that $\psi(x)$ is an analytic function of $|k|$ and extend the original region of $|k| = \sqrt{k_1^2 + k_2^2 + k_3^2}$ $(0 \leq |k| < \infty)$ to a complex k-plane, and consider the analytic continuation of $\psi(x)$ at poles of λ on the upper half k-plane, which we write $\psi'(x)$. Then the argument given below is to show that $\psi'(x)$ satisfies the homogenous integral equation (i. e. $\Psi'(x_1 x_2)$ satisfies Nambu and S. B.'s equation for bound states) and $\psi'(x)$ has the asymptotic form which vanishes exponentially for large values of the relative distance of two particles $|x|$.

$\psi(x)$ is given by*

$$\psi(x)\Big|_{t_1 = t_2} = \frac{1}{(2\pi i)^2} \iint d\epsilon d^3k d^4K \, g\left(\frac{K_0}{2} + k\right) g\left(\frac{K_0}{2} - k\right) \cdot$$

$$\left[\frac{i\left(\frac{K_0}{2} + k\right) + iM}{\left(\frac{K_0}{2} + k\right)^2 + M^2}\right] \cdot \left[\frac{i\left(\frac{K_0}{2} - k\right) + iM}{\left(\frac{K_0}{2} - k\right)^2 + M^2}\right] e^{ikx} \cdot I\theta + \lambda \iint d\epsilon d^3k d^4K \varphi \cdot e^{ikx}.$$

Without loss of generality, we can choose the coordinate system in which the center of mass is at rest. The momentum vector K_0 is then zero. The integration over ϵ in the first term can be performed, considering that mass M has the negative imaginary part $-i\delta$. (δ is an infinitesimal real and positive quantity.) The result is*

$$\frac{-1}{2\pi i} \cdot \int d^3k d^4K \, g\left(\frac{K_0}{2} + k\right) g\left(\frac{K_0}{2} - k\right) \Big\{ \frac{1}{2E - W_0 - i\delta} \cdot \frac{\gamma_4(W_0 - E) - i\gamma k + M}{W_0}$$

$$\cdot \frac{\gamma_4 E + i\gamma k + M}{2E} + \frac{1}{2E + W_0 - i\delta} \cdot \frac{\gamma_4 E + i\gamma k - M}{2E} \cdot \frac{-(W_0 + E)\gamma_4 - i\gamma k - M}{-W_0} \Big\}$$

$$\cdot e^{ikx} I\theta,$$

---

* We assume that $I\theta$ and $g(k)$ have no poles within the region of integration over $|k|$ and ϵ. There is, however, no support to the belief that this assumption is correct. With reference to the definition of $g(k)$, see the footnote on p. 182.

where we put $K_0 = 0$ and $E = \sqrt{k^2 + M^2}$. The second term in this expression concerns with the case that both particles are in negative energy states. Particles with negative energy are, however, treated as the antiparticles with positive energy by reversing the time sense of the world line and the sign of the energy according to Feynman theory.

We introduce polar coordinates for x and k as follows:

$$x = (r \sin \xi \cos \eta, \ r \sin \xi \sin \eta, \ r \cos \xi)$$

$$k = (k \sin \zeta \cos \phi, \ k \sin \zeta \sin \phi, \ k \cos \zeta)$$

and expand $\psi(x)$ in terms of the spherical harmonics $Y_{lm}(\cos \xi, \eta)$.

$$\psi(x)\Big|_{t_1 = t_2} = \sum_{l=0}^{\infty} \sum_{-l \leq m \leq l} \psi_{lm}(r) Y_{lm}(\cos \xi, \eta) . \tag{3.4}$$

Since

$$\exp(ikx) = 4\pi \sum_{l=0}^{\infty} \sum_{-l \leq m \leq l} (i)^l (-1)^m \sqrt{\frac{\pi}{2kr}} J_{l+1/2}(kr) Y_{lm}(\cos \xi, \eta) Y_{l,-m}(\cos \zeta, \phi)$$

and for large r,

$$J_{l+1/2}(kr) \sim (1/2) \sqrt{2/\pi kr} \left[\exp \cdot i \{ kr - (l+1)(\pi/2) \} + \exp -i \{ kr - (l+1)(\pi/2) \} \right],$$

the integration over dk can be carried out after the variable transformation $dE = (\pm)$ $\cdot (2k/E) dk$. By the existence of the δ-factor explicitly written in the denominator, the integration path of E is prescribed to pass below the pole $E = (\pm) W/2$ on the real axis. The result of the integration shows that the arguments k, ϵ_1 and $-k, \epsilon_2$ of $\Gamma \theta$ obey the free particle relation, $k^2 + M^2 = 0$. Hence, using the first equation $(P' \Gamma \theta = \cdot (1-\lambda)\varphi)$ of (2.10), we find

$$\psi_{lm}(r) \sim (2\pi) i^l (-1)^m \iint Y_{l,-m}(\cos \zeta, \phi) d^4 k_1 d^4 k_2$$

$$\left\{ (1-\bar{\lambda})\varphi \frac{e^{i\left(kr - \frac{l+1}{2}\pi\right)}}{kr} + \bar{\lambda}\varphi \frac{1}{kr} \left[e^{i\left(kr - \frac{l+1}{2}\pi\right)} + e^{-i\left(kr - \frac{l+1}{2}\pi\right)} \right] \right\}.$$

Therefore, for positive energy $W_0 *$, we obtain

$$\psi_{lm}(r) \sim (2\pi) i^l (-1)^m \int d\Omega \, Y_{l,-m}(\cos \zeta, \phi) \frac{1}{kr} \left| e^{i\left(kr - \frac{l+1}{2}\pi\right)} + \bar{\lambda} e^{-i\left(kr - \frac{l+1}{2}\pi\right)} \right| \varphi^*$$

$$\cdot (W_0, \cos \zeta, \phi), \tag{3.5}$$

where

$$\varphi(k_1 k_2) = \delta^4(K - K_0) P'(k_1 k_2) \varphi(k_1 k_2),$$

$$\varphi(k_1 k_2) = \left(\frac{-k \cdot M^2}{2\sqrt{k^2 M^2}} \right)^{-1} \varphi^*(W_0, \cos \zeta, \phi).$$

---

* For negative energy states, we also get the same expression, provided that particles with negative energy may be regarded as the antiparticles with positive energy by reversing the time sense of the world line and the sign of the energy.

We consider now the analytic continuation $\psi(r)$ at poles k of λ on the upper half k-plane, which we write $\psi'_{lm}(r)$. The reason why we choose the upper half plane in the complex k-plane is as follows. That is, we have generally

$$k=|\,k\,|=(W_0/2)\,\sqrt{W_0^2-W_m^2}/W_0^2-(Ku)^2,\quad(W_0>0),$$

where

$$W^0{}_m=\sqrt{(2M)^2+K^2},\quad k=|\,k\,|\,u.$$

On account of the square root appearing in the expression for k, we have to made a cut

K χ K

W^{0}_{m}

Fig. 1.

in the W-plane along a suitable curve connecting the singular points $W^0{}_m$ and (Ku) on the real axis in order to make k analytic and one valued throughout the cut plane. As we are interested in the real values of $W_0 < W^0$, we make the cut-curve in the W-plane shown in Fig. 1. And we think that the cut-curve may be chosen as close to the real axis as we like. Putting $a=W_0^{-2}$
$-W^0{}_m{}^2/W_0^2-(Ku)^2$, we define the square root of a by
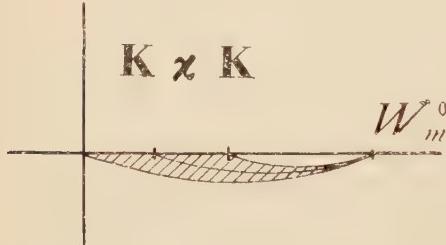
$$\sqrt{a}=\sqrt{\rho}\exp\cdot(i\varphi/2)\quad\text{if}\quad a=\rho\exp(i\varphi)$$

and

$$-\pi+\varepsilon<\varphi<\pi+\varepsilon.$$

Here ε is a finite positive number and may be taken as small as we like. Thus, for the real value of $|\,K\,|<W_0\cdot W^0$, we obtain the positive imaginary part of k in the upper half plane. At poles k of λ on the real axis, $(2\cdot11)$ becomes obviously

$$GI\theta'=\theta'$$

and the asymptotic form $(3\cdot5)$ becomes

$$\psi'_{lm}(r)\sim(2\pi)\,(i)^l(-1)^m(kr)^{-1}\int d\Omega\,Y_{l,-m}(\cos\zeta,\phi)e^{-kr-i\frac{l+1}{2}\pi}\varphi^*.\quad(3\cdot7)$$

The complex eigen-values of the energy W_0 may generally exist corresponding to the value of $k=k'+ik''$.

Further, we shall be able to discuss the same problem as is discussed above starting from the equations $(2\cdot17)$, which leads to the homogeneous equations

$$\left.\begin{aligned}G_2I_{22}\theta_2'+G_2I_{23}\theta_3'=\theta_2'\\G_3I_{32}\theta_2'+G_3I_{33}\theta_3'=\theta_3'\end{aligned}\right\}.\qquad(3\cdot8)$$

The asymptotic behaviours of the wave function and the eigen-values of W_0 will be determined from the poles of λ. But the question how the solutions of $(3\cdot6)$ are related to those of $(3\cdot7)$ is the unsettled problem.

As regards the removal of divergences, we have established the scheme of the renormalization technique in quantum electrodynamics.[10] Though new types of the counter terms are needed to exhibit the possibility of renormalization, (e. g. $-\lambda\phi^4$ for p. s. neutral meson theory,) the considerations given there can apply with a little modifications.

The author should like to express his sincere thanks to Prof. K. Husimi and Prof. R. Utiyama for their helpful discussions. His cordial thanks also go to the members of the research group of Many-Body Problem for their valuable discussions. He is also indebted to Yukawa-Fellowship of Osaka University.

References

1) W. Heisenberg, ZS. f. Phys. **120** (1943), 513, 673
2) F. J. Dyson, Phys. Rev. **75** (1949), 1736.
3) C. Møller, Danske Vidensk. Selskab **23**, no. **1** (1945) ; ibid., no. **19** (1946).
4) H. Kita, Prog. Theor. Phys. **7** (1952), 217.
5) R. P. Feynman, Phys. Rev. **76** (1949), 769.
6) M. Gell-Mann and F. Low, Phys. Rev. **84** (1951), 350.
7) E. E. Salpeter and H. A. Bethe, Phys. Rev. **84** (1951), 1232.
8) J. Hamilton, Proc. Camb. Phil. Soc. **49** (1953), 97.
9) G. C. Wick, Phys. Rev. **80** (1950), 268.
10) S. Nakai, Prog. Theor. Phys. **11** (1954), 155.

Note added in proof : Dr. Hamilton kindly wrote me that his paper[8] contained two mistakes and should be corrected as follows.

i) Equation (4) should be

$$\wedge U \wedge = \wedge \Sigma G \wedge + \wedge \Sigma G P \Sigma G \wedge + \ldots + (1/2) \wedge \Sigma G : \Sigma G \wedge$$
$$+ (1/2) \wedge \Sigma G : \Sigma G P \Sigma G \wedge + (1/2) \wedge \Sigma G P \Sigma G : \Sigma G \wedge + (1/2^2) \wedge \Sigma G : \Sigma G : \Sigma G \wedge + \ldots$$

ii) Equations (5) and (8) are nonsense. Equations (6) and (10) are correct, but they are to be reached by a modified argument.

Dr. Polkinghorne wrote me that he had also spotted these points, corrrected them and extended Dr. Hamilton's work in a different direction.

Analysis of the Structure of Transformation Function
in Quantum Mechanics

Smio TANI

Department of Industrial Chemistry, Faculty of Engineering, Kyoto University

A new device in operator calculus is introduced to treat expressions which contain non-commutable factors. By means of this technique the defining equation for the generator of transformation function of exponential type is derived. It is shown that this equation contains exclusively commutators of relevant quantities.

Then, making use of this peculiarity to the exponential formula, a method is proposed for analyzing the structure of transformation function out of a given Hamiltonian of the system, and this is explained by some examples. This method consists of following procedures: First, decomposition of given Hamiltonian into parts of different character and abstraction of elementary matrices from it. Second, formation of commutators between these matrices and extension, if necessary, of the set of matrices enumerated at the beginning to include new ones which appear through operation of commutation. The matrix ring thus obtained constitutes the frame work of the transformation function.

Through the course of this analysis one can find the ground why problems which are already solved exactly can be treated easily, and also can obtain some instructions how one should proceed in the solution of a given problem.

§ 1. Introduction and Summary

It is frequently necessary to perform a canonical transformation to solve a problem in quantum mechanics. Suppose one must deal with a system with diagonalized Hamiltonian H_0 and the perturbational Hamiltonian H', Then the Schrödinger equation to be solved is

$$(H_0 + H')\Psi = E\Psi , \qquad (1)$$

where E is the eigenvalue of energy of the system. The interaction representation shall be introduced here, since the notation becomes more convenient in treating examples later. If one assumes the adiabatic theorem holds, the considerations in the stationary Schrödinger representation is equivalent to those in the interaction representation.[*] Then one must consider the wave equation

$$i \frac{d\Phi(t)}{dt} = H'(t) \Phi(t), \qquad (2)$$

---

[*] Here we imply by these terms that, if the solution in the interaction representation coincides at the time $t = t_0$ with a certain eigenstate of the unperturbed Hamiltonian, we can eliminate the dependence of this solution on the parameter t_0 by letting $t_0 \to -\infty$, so that we can establish one-to-one correspodence between eigenstates of the unperturbed Hamiltonian and those of total Hamiltonian.

where $H'(t) = \exp(iH_0t) \cdot H' \cdot \exp(-iH_0t)$ and $\varPhi(t)$ is the state vector in the interaction representation and must be imposed a suitable initial condition on it. Of course, H' or $H'(t)$ is not diagonal, but if it were diagonal, one could easily solve eq. (2) by putting the exponential 'ansatz' for $\varPhi(t)$. If the n-th diagonal element of $H'(t)$ in this case is denoted by $\langle n \mid H'(t) \mid n \rangle$, one obtains the n-th component of the solution of the form

$$\langle n \mid \varPhi(t) = \exp\left[-i \int_{t_0}^{t} \langle n \mid H'(\tau) \mid n \rangle\, d\tau \right] \cdot \langle n \mid \varPhi(t_0), \tag{3}$$

where t_0 is the time at which the initial condition is imposed on the system. However, in general the ansatz

$$\exp\left[-i \int_{t_0}^{t} H'(\tau)\, d\tau \right] \cdot \varPhi(t_0)$$

for $\varPhi(t)$ in eq. (2) is not sufficient and it is useful only as the first approximation, except in several cases when the problem is of most simple nature. These simple cases will be discussed in Sec. 3.

Now, it is the purpose of the present paper to point out that the exponential ansatz for the transformation function is most powerful for analyzing its structure. The structure of the transformation function is layed out, although implicitly, as soon as the diagonal and the perturbational Hamiltonians are given concretely. One can get insight into the degree of complexity of the solution, if one succeeds to have some information about the the frame work of this structure. Some necessary informations are offered by the following procedure: First, one decomposes the perturbation Hamiltonian of the interaction representation into several parts of different character by inspection. Practically the result of this decomposition can be written as

$$H'(t) = \sum_n h_n(t)\, M_n . \tag{4}$$

Here M_n stand for matrix operators and h_n for their respective coefficients of ordinary number. (In general h_n are functions of the variable t).

To be specific, let us take the example of the Bloch-Nordsieck type, (a harmonic oscillator with perturbation which is linear in its displacement). In this case

$$H_0 = \omega a^* a, \qquad H' = \lambda \frac{1}{\sqrt{2\omega}}\, (a + a^*), \tag{1'}$$

(ω, the frequency of oscillator: λ, the strength of linear perturbation). And since

$$H'(t) = \frac{\lambda}{\sqrt{2\omega}}\, [(a + a^*) \cos \omega t + i(a - a^*) \sin \omega t], \tag{2'}$$

we choose

$$M_1 = \frac{1}{\sqrt{2}}(a + a^*), \qquad M_2 = \frac{1}{\sqrt{2}}(a - a^*),$$

$$\left(h_1 = \frac{\lambda}{\sqrt{\omega}} \cos \omega t, \qquad h_2 = i\frac{\lambda}{\sqrt{\omega}} \sin \omega t \right). \tag{4'}$$

Then, one forms the commutators between these independent matrices M_n. If new matrices are obtained which are indepent of those which counted at the beginning, one extends the set of matrices so as to include them. This process of commutator formation and extension of the set is continued until the operation of commutator formation becomes closed. *The number of independent elements in the ring thus constructed indicates the degree of complexity of the problem under consideration.*

In the above example, the ring consists of three elements M_1, M_2 and I only. Here I denotes the unit matrix. This is the consequence of the fact

$$[M_1, M_2] = -I.$$

And then the solution of the equation (2) must be written in the form,

$$\Phi(t) = \exp\left[\sum_{m=1}^{N} g_m(tt_0)\overline{M}_m\right] \cdot \exp\left[iW(t_0-t)\right] \cdot \Phi(t_0), \tag{5}$$

where N is the number of the independent elements of the ring mentioned above and $\overline{M}_m$ denotes any one of its elements. And W in the second member on the right hand side is the operator of the level shift caused by the perturbation H', and its meaning will be fully explained later. It is possible to choose W diagonal. In view of the initial condition imposed at $t=t_0$, the functions g_m must satisfy the conditions $g_m(t_0 t_0) = 0$. This statement is concluded from the fact that if one puts the exponential ansatz for the transformation function and starts from

$$\Phi(t) = \exp\left[iG(t)\right] \cdot \chi(t), \tag{6}$$

then, by eq. (2) and through some manipulations, the following relation is obtained. The operator $G(t)$ should be defined by the relation,

$$\frac{dG}{dt} = -H' - \frac{1}{2}\left[H', iG\right] - \sum_{n=1}^{\infty} \frac{(-1)^n B_n}{(2n)!}\left[\cdots\left[H', \overbrace{iG\right]\cdots iG}^{2n-\text{times}}\right]$$

$$+ W - \frac{1}{2}\left[W, iG\right] + \sum_{n=1}^{\infty} \frac{(-1)^n B_n}{(2n)!}\left[\cdots\left[W, \overbrace{iG\right]\cdots iG}^{2n-\text{times}}\right], \tag{7}$$

where B_n is the Bernoulli number and W is the operator of the level shift mentioned above. And also one obtains the wave equation in the new representation obtained through the transformation generated by e^{iG}, that is,

$$i\frac{d\chi}{dt} = W\chi. \tag{8}$$

(It is feasible to solve eq. (7) and at the same time to determine the diagonal operator W, if once the Hamiltonian H' is given. But this point will be discussed in the separate paper in which the adiatic theorem is investigated.) On observing the right hand side of eq. (8), it is noticeable that it consists exclusively of commutators. As will be shown

later, the existence of terms containing W causes no difficulty ; and the reader may neglect them for a moment. Consequently it is sufficient to put

$$i G(t) = \sum_{m=1}^{N} g_m(tt_0) \cdot \bar{M}_m \tag{9}$$

for $G(t)$ and compare the coefficients of each matrix $\bar{M}_m$ on both sides of eq. (7), since the matrix ring is closed as regards to mutual commutation. On this ground it can be understood naturally that the above mentioned procedure of analyzing the structure of transformation function, i. e. construction of the matrix ring $\{\bar{M}_m\}$ and counting the number of its elements, is reasonable. In the next section we shall prove that appearence of commutators only is peculiar to the exponential transformation function. This is the reason why we insist on employing it in the analysis of the structure of transformation function.

However, the superiority of the exponential ansatz in the case of exploring the frame work of the transformation function does not necessarily mean that the practical calculation can be carried out most conveniently in terms of the exponential transformation function. It may happen according to the type of a problem that an ansatz of other type than the exponential is more convenient to deal with. For example, in some cases it is known that the half-arctangent ansatz $\exp|(i/2)\tan^{-1}G|$ is most convenient, since the function G in such cases are simple. The Cayley transformation is also useful in several cases. In the following we generally call such an operator G as *the generator of the transformation function* with which the transformation is expressed and which we expect to make manipulations most convenient. The characteristics of these kinds of transformation function will be discussed in the last section.

The proof of the formula (7) is carried out by means of a new artifice for arranging noncommutable factors. This artifice consists of rather trivial devices of operator calculus, but has been employed in several problems* and proved useful in performing several kinds of canonical transformation. As will be seen later, the fact that eq. (7) contains exclusively commutators will become evident, as soon as one makes ure of this method.

Section 2 is concerned with remarks on the technique of treating noncommutable factors mentioned just above and on the derivation of the formula for the exponential transformation function. In Sec. 3 the method for the analysis of the structure of transformation function is explained, and an example of somewhat different kind from that referred above is dealt with. This example has been stimulated by the study of nuclear forces. It is concerned with a question ' Are the wave functions in the coupled $^3S + ^3D$ state of $n-p$ system with constant potentials (central and tensor) is different in character** from those with variable potentials (for example, the Yukawa potential), or not ? '. The method

---

* In separate papers, we shall treat the problem concerning Furry's bound state representation and the meaning of one-body treatment when applied to a quantized fermion field, and the other problem concering elimination of pair effects and derivation of effective Hamiltonian in $Ps - Ps$ theory.

** If the centrifugal force were absent in the D state, then the wave function of the system with constant potentials become so simple that we may differentiate it from the general ones with variable potentials. Our question is ' What is the situation in the presense of the centrifuqal force ?'.

proposed in this paper is applied, and the answer is 'No'. In section 4 there are given several remarks in carrying out canonical transformations in quantum mechanics.

§ 2. New technique in treating noncommutable factors and derivation of the formula for exponential transformation function

We must prove eq. (7) of the previous section. Before the proof, it is desirable to introduce a new artifice in arranging noncommutable factors ; so let us begin with giving some remarks on the treatment of noncommutable factors.

First, remind ourselves how one usually proceeds in solving the wave equation (2) by means of the exponential ansatz (6). The following formula is commonly known

$$e^{-G} A \, e^{G} = A + [A, G] + \frac{1}{2!} [[A, G], G] + \cdots\cdots, \qquad (10)$$

where G and A may dedote any two operators. This result can be shown directly if one adopts the following method of operator calculus ; and the method is generally applicable in treating any operator of the form

$$f(G) A \, g(G) \qquad (11)$$

where both f and g are functions which are defined by series of non-negative power in G. The operators denoted by $\overrightarrow{G}$ and $\overleftarrow{G}$ are introduced, and they multiply the operand by G from the left and from the right, respectively. Then regarding the factor A as the operand to be operated by these, one writes

and

$$\overrightarrow{G} \cdot A \qquad \text{for} \qquad G A$$
$$\overleftarrow{G} \cdot A \qquad \text{for} \qquad A G, \qquad \text{respectively.} \qquad (12)$$

Here one puts the dot to separate the operators in the new sense from their operand.

Now, it is often desirable to rearrange the factors in an expression like (11) in order to reduce the whole expression to a form as concise as possible. Each time a rearrangement of factors is made, one has dealt with some commutators, so that it is convenient to introduce here a special notation for the operator of commutator formation ; and we write

$$C_{G} \cdot A \qquad \text{for} \qquad [A, G] \qquad (13)$$

of course, the operator C_{G} is related to $\overrightarrow{G}$ and $\overleftarrow{G}$ by

$$C_{G} = \overleftarrow{G} - \overrightarrow{G}. \qquad (14)$$

A multiple commutator (e. g. of degree n)

$$[\cdots[\overbrace{A, G], \cdots, G}^{n \text{ times}}]$$

---

* Analogous considerations were once made by Prof. Y. Nambu several years ago for somewhat different purposes from those here discussed. Cf. Ref. 1.

can be written as $C_G^n \cdot A$; the consistency of this notation will be explained below. Thus one may write

$$f(\overrightarrow{G})g(\overleftarrow{G}) \cdot A \text{ or } f(\overrightarrow{G})g(\overrightarrow{G}+C_G) \cdot A \text{ and so on for } f(G)Ag(G). \quad (11')$$

One should notice to the fact that operators $\overrightarrow{G}$, $\overleftarrow{G}$ and C_G are commutable with each other by their definition. Consequently, one may treat any of their function on the left side of the dot as if they were ordinary numbers. This fact makes it very easy to foresee the result of rearrangement performed in expressions like (11). It is evident that the relation (10) holds, since

$$e^{-G} A e^{G} \equiv (e^{-\overrightarrow{G}} e^{\overleftarrow{G}}) \cdot A = e^{\overleftarrow{G}-\overrightarrow{G}} \cdot A = e^{C_G} \cdot A \ . \quad (15)$$

The consistency of the method now introduced can be concluded as follows. One can prove the following identity by means of the induction ;

$$AG^n = G^n A + nG^{n-1}[A, G] + \cdots + \frac{n!}{m!(n-m)!} G^{n-m}[\cdots[\overbrace{A, G]\cdots, G}^{m-\text{times}}] + \cdots. \quad (16)$$

On the other had, one has equally the same result according to the new method ; it is an immediate consequence of the relation (14), that is, one has

$$\overrightarrow{G^n} = (\overrightarrow{G}+C_G)^n = \sum_{m=0}^{n} \frac{n!}{m!(n-m)!} \overrightarrow{G}^{n-m} C_G^m . \quad (17)$$

Moreover, as $\overrightarrow{G}$ and $\overleftarrow{G}$ are commutable with each other by their definition, the higher powers of C_G are given unambiguous meaning to them, and one obtains the result written out in (16) by applying both hand sides of eq. (17) on the operand A. It is valid so long as functions defined in terms of non-negative power series in G are concerned.

Next, let us go on reflecting how one usually performs a canonical transformation with the exponential transformation function. Suppose for a moment that the perturbation Hamiltonian is proportional to a small parameter, denoted by λ. Then one may assume that the generator G can be expanded in powers of λ, and puts it into the form

$$G = \lambda G_1 + \lambda^2 G_2 + \cdots . \quad (18)$$

The substitution of (6) into (2) yields the equation to be solved

$$\left(e^{-iG} i \frac{d}{dt} e^{iG} \right) \chi = (e^{-iG} H' e^{iG}) \chi,*$$

that is,

---

* Here d/dt is considered as a kind of matrix operator, so that dG/dt is the substitute for $[d/dt, G]$. The operator d/dt differentiates all factors standing on its right hand side ; χ is differentiated inspite of the bracket which stands between it and d/dt.

$$i\frac{d\chi}{dt} = \left\{ H' + [H'iG] + \cdots + \frac{dG}{dt} + \frac{1}{2!}\left[\frac{dG}{dt}, iG\right] + \cdots \right\}\chi . \qquad (19)$$

Herefrom one obtains the defining equation for G_1, which is of first order in λ,

$$\lambda\frac{dG_1}{dt} + H' = 0 . \qquad (20)$$

H' is nondiagonal and must be eliminated from eq. (19) by a suitable definition of G_1. The operator $[H', iG_1]$ may contain both diagonal and nondiagonal parts; these are denoted by $\langle [H', iG_1]\rangle_D$ and $\langle [H', iG_1]\rangle_{ND}$, respectively. Then G_2 must be defined so that nondiagonal terms of second order in λ may be eliminated, and one has the defining equation for G_2,

$$\lambda^2\frac{dG_2}{dt} + \frac{1}{2}\langle [H', i\lambda G_1]\rangle_{ND} = 0 . \qquad (21)$$

Higher order terms may be defined similarly. And one finally finds the wave equation after the transformation generated by e^{iG} as

$$i\frac{d\chi}{dt} = \left\{ \frac{\lambda}{2}\langle [H', i\lambda G_1]\rangle_D + \cdots \right\}\chi . \qquad (22)$$

Here $(\lambda/2)\langle [H', iG_1]\rangle_D$ etc. represent the operators of the level shift caused by the perturbation H'.

Finally, we generalize these considerations and derive the general formula (7) for defining the generator G. Our task is to obtain from (1), (6) and (8), the transformed wave equation of the form

$$\left\{ e^{-iG}\left(\frac{1}{i}\frac{d}{dt} + H'\right)e^{iG}\right\}\chi = \left\{\frac{1}{i}\frac{d}{dt} + W\right\}\chi = 0 ,$$

where W is the level shift which is the generalization of such factors as $(\lambda/2)$ $\cdot\langle [H', iG_1]\rangle_D$ in eq. (22). According to the method introduced above we have

$$(e^{-iG} e^{iG})\cdot\left(\frac{1}{i}\frac{d}{dt} + H'\right) = 1\cdot\left(\frac{1}{i}\frac{d}{dt} + W\right), \qquad (23')$$

that is,

$$(e^{iG} - 1)\cdot\frac{1}{i}\frac{d}{dt} = 1\cdot W - e^{iG}\cdot H' . \qquad (24)$$

Then we obtain the result (7), since we have

$$C_G\cdot\frac{d}{dt} = \frac{iC_G}{e^{iG} - 1}\cdot W - \frac{iC_G e^{iG}}{e^{iG} - 1}\cdot H' . \qquad (25)$$

Eq. (7) is the generalization of such defining equations as one meets in eqs. (20) and

(21). Thus our proof is completed.* The fact that eq. (7) contains exclusively com-mutators is an immediate consequence of the fact that there appears C_G only in eq. (23′) or (24) ; and as can be easily shown, this is peculiar to the exponential transformation function.

§ 3. Analysis of the structure of transformation function

In this section, we shall investigate the structure of the transformation function. Thereby the formula (7) for definig the generator G will be taken as the basis for all processes of analysis.

At the beginning let us study examples which are solved by the canonical transforma-tion of relatively simple structure. These examples have been discussed by several authors,[3] and some of them are utilized in many physical problems. We shall begin by enumerating cases of simplest type, and then explain how they are easily solved by assuming the exponential transformation function.

(i) First, the simplest case among others is the one in which $H'(t)$ is already diagonal. This case is rather trivial, and we have nothing to add to the remark at the begining of Sec. 1.

(ii) Second is the case in which all matrix elements of $H'(t)$ have the same depe-dence on t, i. e. $H'(t)$ is a constant matrix times a certain function of t.

(iii) Third is the case, in which the matrix $H'(t)$ is singular, so that its product by its time-integral** vanishes, i. e.

$$H'(t) \int^t H'(t) = 0 . \qquad (26)$$

In these cases, if one assumes the exponential ansatz for the transformation function, then the defining equation (7) for the generator reduces to the simplest form ; that is, only the first term of the first line in (7) is left and all other terms vanish,

$$\frac{dG}{dt} = -H'(t), \qquad (27a)$$

and one may put

$$W = 0 , \qquad (28a)$$

and the solution of eq. (2) is given by

$$\Phi(t) = \exp\left[\frac{1}{i} \int_{t_0}^t H'(\tau) d\tau \right] \cdot \Phi(t_0). \qquad (29a)$$

---

* The result given in eq. (7) is suggested in the old paper by Hausdorf,[2] who treated some analogous equations. But he computed in an elementary way, that is, he formed multiple commutators of the original equation with quantities like G and then computed their linear combinations so as to cancel terms which contain commutators of dG/dt. The author is very much indebted to Prof. K. Husimi for sending him the note of Hausdorf's paper several years ago.

** By the symbol $\int^t dt$, it is meant that the indefinite integral should be taken.

(iv)   The next case, which is less simple than cases (i)—(iii), is the one referred to in Sec. 1 as the example, and the processes of constructing the matrix ring and counting its elements is performed there.   That was a harmanic oscillator with perturbation linear in its displacement.   In general terms, in those cases to be classified into the *genre* now considered, the given Haimiltonian $H'(t)$ has such a property that the commutator of $H'(t)$ with its time-integtal reduces to a unit matrix times an ordinary number,

$$\left[H'(t), \int^t H(t)dt \right] = c(t) \cdot I . \tag{30}$$

Then eq. (7) for the generator of exponential type reduces to first three terms, (two terms in the first line and the term W), only.

$$\frac{dG}{dt} = -H(t) - \frac{1}{2}[H'(t), iG] + W, \tag{27b}$$

beause in this case the higher multiple commutors in the right hand side vanish on account of the condition (30).   The solution is given by

$$\Phi(t) = \exp\left[\frac{1}{i} \int^t H'(t)dt \right] \cdot \chi(t), \tag{29b}$$

where new state function $\chi(t)$ must be given by solving the diagonalized equation

$$-\frac{d\chi}{dt} = \frac{1}{2}\left[H'(t), \int^t H'(t)dt \right] \cdot \chi(t), \tag{31}$$

under the corresponding initial condition.

In examples (i)—(iv) the solution is obtainable by means of successive approximation. If we imagine the interaction Hamiltonian is proportional to some small parameter, λ, then after expanding both sides of eq. (7) into parts proportional to each power in λ. G can be obtained by integrating the corresponding member on the right hand side.   As regards to the first few terms the result is

$$G(t) = -\int^t H'(t)dt + \frac{i}{2}\int^t \left\langle\left[H'(t), \int^t H'(t)dt \right]\right\rangle_{NN}dt + \cdots . \tag{32}$$

Actually in these examples of simple nature, the processes of successive approximation come to an end, and one obtains such solutions as described in (29a) and (29b).   But since these solutions are valid irrespective of the magnitude of λ, one thus reaches to the answer of these problems by means of the exponential transformation function.

There are known problems of a little more complicated nature which are solved exactly by means of a canonical transformation.   Two of them shall be mentioned ; (i) a harmonic oscillator with perturbation which is quadratic in its canonical variables, and (ii) a rotator in a three-dimensional space which is perturbed by a torque applied in the direction perpendicular to the original axis of rotation.

As regards to the example of the first type, it is sufficient to consider the case in which the perturbation is bilinear in x and p, that is, the case with

$$H_0 = \omega a^* a, \quad H' = \frac{\lambda}{2}(xp + px) = \frac{\lambda}{2i}(a^2 - a^{*2}), \tag{33a}$$

because the cases with perturbation of the form λx^2 or λp^2 can be manipulated without difficulty by suitable renormalization of the frequency and the amplitude of the oscillator. Also the general case with perturbation of the form $(\lambda/2)(xp + px) + \lambda' x^2 + \lambda'' p^2$ can be reduced to the above type by a similar way. In this case the Hamiltonian in the interaction representation is given as

$$H'(t) = \frac{\lambda}{2}\left[\left(\frac{a^2 - a^{*2}}{i}\right)\cos 2\omega t + (a^2 + a^{*2})\sin 2\omega t\right]. \tag{34a}$$

As for the example of the second type, we choose the case of 1/2-spin rotator. And the Hamiltonian in this case may be taken without losing generality as

$$H_0 = \omega \sigma_3, \quad H' = \lambda \sigma_1, \tag{33b}$$

then after going over to the interaction representation, we have

$$H'(t) = \lambda[\sigma_1 \cos 2\omega t - \sigma_2 \sin 2\omega t]. \tag{34b}$$

In these cases eq. (7) from which we start cannot be cut by a certain finite number of terms. And the process of successive approximation must be continued to infinite order. However, the fact that eq. (7) contains exclusively commutators suggests us what is the form of the solution. There are two types of matrix operators appearing in the Hamiltonian of the interaction representation. Let us denote them as M_1 and M_2. If we form the commutator $[M_1, M_2]$, then we obtain a new matrix operator, denoted by M_3. M_3 is diagonal, though not a unit matrix times a constant as in the last example (iv). By further forming the commutators of M_3 with others, we find that the matrices M_1, M_2 and M_3 form a closed set.

Specifically, in the first example we have

$$M_1 = \frac{a^2 - a^{*2}}{2}, \quad M_2 = \frac{a^2 + a^{*2}}{2}, \quad M_3 = aa^* + a^*a, \tag{35a}$$

and in the second example we have

$$M_1 = \sigma_1, \quad M_2 = \sigma_2, \quad M_3 = 2i\sigma_3. \tag{35b}$$

Through above considerations we are lead to the conclusion that the generator G should be of the form

$$G = g_1(t) M_1 + g_2(t) M_2, \tag{36}$$

and the operator of the level shift which is diagonal should be of the form

$$W = g_3(t) M_3, \tag{38}$$

where g_1, g_2 and g_3 are functions to be chosen so that eq. (7) may be satisfied. The correctness of these statements can be easily checked by substitution of them into (7). The practical computation of G and W may be carried out by starting from eq. (7) and by summing an infinite series. But it is also possible to choose another type of transformation function than the exponential, for example, the type of half-arctangent, $\exp[(i/2) \cdot \tan^{-1} G]$, and then to reduce the generator G of the new form into a simple function of parameters given in the Hamiltonian under consideration. This point will be discussed

in the next section.

The considerations in the above example show how one should proceed in a more general case. One must first decompose the Hamiltonian in the interaction representation into parts of different character and abstract elementary matrix operators from it. Next, one must form commutators between these matrices and extend the set of matrices so as to include the new matrices thus formed. The closed set of matrices constructed in this way constitutes the frame work of the transformation function. The number of elements in this set indicates the degree of complexity of the transformation. *It gives the minimum number of elements which are necessary in constructing the solution of given problem.* The transformation function should have the form

$$G(t) = \sum_n{}' g_n(t) \cdot M_n^{ND},$$ (38)

and the operator of level shift should have the form

$$W(t) = \sum_n{}' W_n(t) \cdot M_n^D,$$ (39)

where M_n^{ND} and M_n^D denote the members of the matrix ring, which are nondiagonal and diagonal respectively.

Finally we add another type of example to show that our analysis actually gives minimum number of elements which are necessary. As mentioned at the end of Sec. 1, this example is stimulated by some practical interests in the study of nuclear force. In a phenomenological treatment of the ${}^3S + {}^3D$ state of $n-p$ system, one may raise a question whether the use of square well in both central and tensor forces make the wave functions simpler than those computed using potentials which are variable with distance so much as to make the former preferable in obtaining qualitative aspects of the problem. This is surely so if the centrifugal potential in D state were absent. But when the centrifugal force is taken into account, actual computation of wave functions is troublesome and one must resort to means such as the numerical integration, even when square wells are adopted for central and tensor forces. However, it may happen that the case with square well is simpler than others with variable potentials in the sense that some approximation, for example the W. K. B.-method, is suitable in the one case while it is not so in the others. This question is motivated by inspection of the equation to be solved : As will be explained in the Appendix I, the coupled second order differential equations can be written in a linearized form[4] and it is of just the same form as eq. (2) except that imaginary unit i disappears from the left hand side. If the 'Hamiltonian' $H(x)$ in our case is decomposed into parts with different functional dependence on x, we find that the number of elementary matrices is *two in the case with square wells* and *four in the case with variable wells*. Then, according to the prescription discussed, the analysis is performed in order to answer the raised question. By forming commutators, starting with two or four matrices in each case, and by extending the set of matrices, we finally find *ten* independent matrices to be taken into account in both cases. Thus we conclude that there should be no difference between square well and variable well as

regards to the character of the solution calculated with them. It is noticable only ten matrices are sufficient to be considered; because as the Hamiltonian in our example is a 4—4-matrix, the maximum number of elements to be considered comes to *sixteen*. But some of sixteen elements appear into our solution always as certain combinations of them, so that only ten matrices suitably chosen serve as the minimum number of matrices which is necessary in the solution. The process of commutator formation is most suitable in exhibiting these necessary elements.

§ 4. Discussions

In the previous sections it is shown that the exponential ansatz for the transformtion function is most suitable for the analysis of its structure. One can point out two grounds for this statement. (i) As discussed in Sec. 2, the defining eq. (7) for the generator of exponential type contains only commutators of relevant quantities. (ii) As discused in Sec. 3, the processes of commutator formation enable one to obtain minimum number of necessary matrices in writing down the solution.

However, as already mentioned, the above situation does not necessarily mean that the exponential ansatz is in every case best for practical calculation of the transformation function. For example, when we treat with a rotation in a 1/2-spin space, it happens that the most suitable ansatz for the transformation function is to put it in the form

$$U = e^{t/2 \tan^{-1} G} = \left(\frac{1+iG}{1-iG} \right)^{1/4} \tag{40}$$

and look for a certain Hermitian operator G. This is related to the fact that every rotation in this space may be generated by means of a transformation of the form

$$U = \exp\left[\frac{i}{2} (\sigma u) \varphi \right] \tag{41}$$

(where σ denotes spin matrices, and u denotes a unit vector in this space which defines the axis of rotation, and φ is the angle of rotation), and that one often finds a simple expression for $\tan \varphi$ through inspection of the problem. The transformation which diagonalizes the Hamiltonian of a free Dirac particle[5] and the Foldy transformation frequently used in pseudoscalar meson theory with ps-coupling[6] are two examples of this kind. The defining equation for the generator of half-arctangent ansatz is given as

$$2\frac{dG}{dt} = -H' - i[H', G] - G iH'G - (1+G^2)^{1/2} H' (1+G^2)^{1/2}$$
$$- (1+G^2)^{1/4}[(1-iG)^{1/2} H'(1+iG)^{1/2} + (1+iG)^{1/2} H'(1-iG)^{1/2}](1+G^2)^{1/4}$$
$$+ W - i[WG] + GWG + (1+G^2)^{1/2} W(1+G^2)^{1/2} \tag{42}$$
$$+ (1+G^2)^{1/4}[(1-iG)^{1/2} W(1+iG)^{1/2} + (1+iG)^{1/2} W(1-iG)^{1/2}](1+G^2)^{1/4}.$$

This result is obtained in Appendix II by means of the technique proposed in Sec. 2 In

such examples as referred to in the above, the higher order terms on the right hand side vanish and the answer is given simply by

$$G = -2 \int^{t'} H'(t) dt , \tag{43}$$

with the diagonal operator of level shift which is defined by

$$W = \frac{-iG}{1 + (1 + G^2)^{1/2}} H' . \tag{44}$$

Another example is the transformation of Cayley type

$$U = e^{\pi i \tan^{-1} G} = \frac{1 + iG}{1 - iG} , \tag{45}$$

which is most suitable when we must proceed by means of power series expansion in order to determine the transformation function. This is the case when the higher order perturbational terms caused by the nondiagonal part of the original Hamiltonian is much complicated and we have no better method. The usefulness of this type of transformation function is related to the fact that it yields the defining equation for the generator G which has simplest coefficients among others. The defining equation in this case can be written as

$$2\frac{dG}{dt} = -H' - i[H'G] - GH'G + W - i[WG] + GWG, \tag{47}$$

which is shown in Appendix II.

In any case the most suitable type of ansatz depends upon the type of problem, but for a qualitative analysis of the structure of transformation the exponential ansatz is sufficient and convenient among others.

Finally we should like to give a remark on the artifice of rearranging noncommutable factors, which is proposed in Sec. 2. There our consideration is restricted to functions of non-negative power in G. The generalization to include negative powers may be possible. We obtain at a certain integer n the following result

$$G^{-n}A = G^{-n}AG^{n}G^{-n} = \sum_{m=0}^{n} \frac{n!}{m!(n-m)!} G^{-m} \overbrace{[\cdots[AG]\cdots G]}^{m-\text{times}} G^{-n}, \tag{47}$$

as the consequence of eq. (16) which is established by the induction. And by repeated application of this formula we have the formal expansion of the results which is written as follows,

$$G^{-n}A = AG^{-n} + n[AG]G^{-(n+1)} + \frac{n(n+1)}{2}[[A, G]G]G^{-(n+2)} + \cdots . \tag{48}$$

But this expression can be deduced from the application of the formal expansion of the following operator into powers of $(\overleftarrow{C_G/G})$

$$\overrightarrow{(G)}^{-n}=(\overrightarrow{G}-C_G)^{-n}=(\overleftarrow{G})^{-n}[1-(C_G/\overleftarrow{G})]^{-n}=(\overleftarrow{G})^{-n}+nC_G(\overleftarrow{G})^{-(n+1)}+\cdots,$$

$$(49)$$

and application of both sides on the operand A. However, some consideration concerning the convergence of the series should be necessary for the consistency of this generalization.

The process of constructing a matrix ring through commutator formation may suggest some connection of our method to the theory of Lie group, since in this theory Lie ring is constructed in a similar way. But we do not go into this aspect of the problem.

In this paper only discussions of general nature are given. Applications to physical problems will be found in subsequent papers. In view of the importance of performing canonical transformations in order to treat systems with not small perturbation, a paper of nature like the present one may be allowed to find its place.

Appendex 1. *On the differential equations used in $^3S+{}^3D$ state of n-p system*

The coupled second order differential equations are of the form (in a suitable unit)

$$\left.\begin{array}{l} u''+A(x)u+C(x)v=Eu, \\ v''+C(x)u+B(x)v=Ev. \end{array}\right\} \qquad (A\cdot1)$$

$$\left.\begin{array}{l} A(x)=V_C(x), \\ B(x)=V_C(x)-2V_T(x)-\dfrac{6}{x^2}, \\ C(x)=2^{3/2}V_T(x), \end{array}\right\} \qquad (A\cdot2)$$

where u and v denotes the wave function of S and D state respectively. V_C and V_T is the potential of central and tensor force respectively. Now, these equtions are linearized[1] through introduction of the ' Hamiltonian ' $H(x)$ and the state function $\Psi(x)$. For this purpose we should put

$$\frac{d\Psi(x)}{dx}=H(x)\Psi(x), \qquad (A\cdot3)$$

$$\Psi(x)=\begin{pmatrix} \dot{u} \\ u \\ \dot{v} \\ v \end{pmatrix}, \qquad H(x)=\begin{pmatrix} 0 & E-A(x) & 0 & -C(x) \\ 1 & 0 & 0 & 0 \\ 0 & -C(x) & 0 & E-B(x) \\ 0 & 0 & 1 & 0 \end{pmatrix}. \qquad (A\cdot4)$$

Thus in the case of square wells we have two elementary matrices appeared in the

Hamiltonian.

$$M_1 \equiv \begin{pmatrix} 0 & C_1 & 0 & C_3 \\ 1 & 0 & 0 & 0 \\ 0 & C_2 & 0 & C_4 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \qquad M_2 \equiv \begin{pmatrix} 0 & \cdot & \cdot & 0 \\ 0 & \cdot & \cdot & 0 \\ 0 & 0 & 0 & 1 \\ 0 & \cdot & \cdot & 0 \end{pmatrix}, \qquad (A\cdot 5)$$

where C_1, C_2, C_3 are certain constants, because only $B(x)$ depends on x through centrifugal potential, $6/x^2$. The processes of commutator formation exhibits the set of independent matrices necessary in writing out the solution ; i. e.,

$$M_1, \ M_2, \ [M_2 M_1], \ [M_2 M_1 M_1], \ [M_2 M_1 M_1 M_1],$$

$$[M_2 M_1 M_1 M_1 M_1], \ [M_2 M_1 M_1 M_1 M_1 M_1], \ [M_2 M_1 M_1 M_1 M_2], \qquad (A\cdot 6)$$

$$[M_2 M_1 M_1 M_1 M_1 M_2], \ [M_2 M_1 M_1 M_1 M_1 M_2 M_1],$$

(multiple commutators are abbreviated as $[M_1 M_2 M_3]$ for $[[M_1 M_2], M_3]$). On the other hand, in the case of variable wells four elementary matrices appear in the Hamiltonian

$$M_1 \equiv \begin{pmatrix} 0 & \cdot & \cdot & 0 \\ 1 & 0 & 0 & 0 \\ 0 & \cdot & \cdot & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \ M_2 \equiv \begin{pmatrix} 0 & 1 & 0 & 0 \\ \cdot & \cdot & 0 & 0 \\ \cdot & \cdot & 0 & 0 \\ \cdot & \cdot & 0 & 0 \end{pmatrix}, \ M_3 \equiv \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & \cdot & \cdot & 0 \\ 0 & 1 & 0 & 0 \\ 0 & \cdot & \cdot & 0 \end{pmatrix}, \ M_4 \equiv \begin{pmatrix} 0 & \cdot & \cdot & 0 \\ 0 & \cdot & \cdot & 0 \\ 0 & 0 & 0 & 1 \\ 0 & \cdot & \cdot & 0 \end{pmatrix}.$$

$$(A\cdot 7)$$

And the set of independent matrices consists of elements

$$M_1, \ M_2, \ M_3, \ M_4, \ [M_1 M_2],$$

$$[M_1 M_3], \ [M_1 M_4], \ [M_1 M_4 M_1], \ [M_1 M_2 M_1], \ [M_1 M_3 M_1 M_2]. \qquad (A\cdot 8)$$

It can be shown that the set $(A\cdot 6)$ is equivalent to the set $(A\cdot 8)$.

Appendix II. *On transformation of half-arctangent type and Cayley type*

When we put the half-arctangent ansatz for the transformation function, the defining equation should be derived from

$$\left(\frac{1-iG}{1+iG}\right)^{1/4}\left(\frac{1}{i}\frac{d}{dt} + H'\right)\left(\frac{1+iG}{1-iG}\right)^{1/4} = \frac{1}{i}\frac{d}{dt} + W \qquad (A\cdot 9)$$

(corresponding to eq. (23) for the exponential ansatz). Then we have

$$\left[\left(\frac{1-i\overrightarrow{G}}{1-i\overrightarrow{G}}\right)^{1/4}\left(\frac{1+i\overleftarrow{G}}{1-i\overleftarrow{G}}\right)^{1/4} - 1\right]\cdot\frac{1}{i}\frac{d}{dt} = 1\cdot W - \left[\left(\frac{1-i\overrightarrow{G}}{1+i\overrightarrow{G}}\right)\left(\frac{1+i\overleftarrow{G}}{1-i\overleftarrow{G}}\right)\right]^{1/4}\cdot H',$$

$$(A\cdot 10)$$

which can be rewritten as

$$\frac{1}{i}(\overleftarrow{G}-\overrightarrow{G})\cdot\frac{d}{dt}$$

$$=-\frac{i(\overleftarrow{G}-\overrightarrow{G})}{(1+i\overrightarrow{G})^{1/4}(1-i\overrightarrow{G})^{1/4}-(1-i\overrightarrow{G})^{1/4}(1+i\overleftarrow{G})^{1/4}}$$

$$\cdot\{(1+i\overrightarrow{G})^{1/4}(1-i\overrightarrow{G})^{1/4}\cdot W-(1-i\overrightarrow{G})^{1/4}(1+i\overrightarrow{G})^{1/4}\cdot H'\}, \quad (A\cdot11)$$

and then we have the result (42) in the text after some algebraic manipulation.

If we treat with a $1/2$-spin space, our equation for the half-arctangent transformation function is able to be reduced further. In this case spin matrices have the property

$$\sigma_i^2=1, \ i=1, 2, 3 .$$

And consequently, we can classify the unit matrix and σ_3, which are taken as diagonal, into one group and call them even, while nondiagonal matrices σ_1 and σ_2 are called odd. Then, we should take odd operators for the generator G, and even ones for the operator of level shift W. (This is in accordance with the analysis made in Sec. 3, eq. (33b) and forth, because then $\tan^{-1}G$ is odd and have required structure). Comparing even parts of both sides of eq. (42) we have

$$0=W+GWG+(1+G^2)^{1/2}W(1+G^2)^{1/2}$$

$$+(1+G^2)^{1/4}[(1-iG)^{1/2}W(1+iG)^{1/2}+(1+iG)^{1/2}W(1-iG)^{1/2}](1+G^2)^{1/4}$$

$$+i[G, H'],\qquad\qquad\qquad (A\cdot12)$$

from which W can be derived, and the result is, by making use of the operator calculus again,

$$W=\{1+\overrightarrow{G}\,\overleftarrow{G}+(1+\overleftarrow{G^2})^{1/2}(1+\overrightarrow{G^2})^{1/2}+(1+\overleftarrow{G^2})^{1/4}(1+\overrightarrow{G^2})^{1/4}$$

$$\cdot[(1-i\overrightarrow{G})^{1/2}(1+i\overleftarrow{G})^{1/2}+(1+i\overrightarrow{G})^{1/2}(1-i\overleftarrow{G})^{1/2}]\}^{-1}\cdot i[H', G]. \ (A\cdot13)$$

Thus W is eliminated from defining equation for G, and we have

$$\frac{dG}{dt}=-2H'-\frac{1}{2}(G^2H'+2GH'G+H'G^2)+\frac{1}{2}[[H',(1+G^2)^{1/2}],(1+G^2)^{1/2}]$$

$$(A\cdot14)$$

from which we obtain the simple results for G, eq. (43), in examples mentioned in the text. As the results of the fact

$$H'(t)\int^t H'(t)dt+\int^t H'(t)dt\,H'(t)=0,$$

much simplification is attained, and this fact also simplifies the expression for W, eq. (44),

which is easily derived from $(A \cdot 13)$ using the fact that $\boldsymbol{G}\boldsymbol{H}' + \boldsymbol{H}'\boldsymbol{G} = 0$. i. e., $\overrightarrow{\boldsymbol{G}} + \overleftarrow{\boldsymbol{G}} = 0$ when operated on $\boldsymbol{H}'$.

When we put the Cayley ansatz for the transformation function, our starting point is

$$\left(\frac{1-i\boldsymbol{G}}{1+i\boldsymbol{G}}\right)\left(\frac{1}{i}\frac{d}{dt} + \boldsymbol{H}'\right)\left(\frac{1+i\boldsymbol{G}}{1-i\boldsymbol{G}}\right) = \frac{1}{i}\frac{d}{dt} + \boldsymbol{W}, \qquad (A \cdot 15)$$

(corresponding to eqs. (23) or $(A, 6)$ in the other cases). On rewriting it as

$$\left[\left(\frac{1-i\overrightarrow{\boldsymbol{G}}}{1+i\overrightarrow{\boldsymbol{G}}}\right)\left(\frac{1+i\overleftarrow{\boldsymbol{G}}}{1-i\overleftarrow{\boldsymbol{G}}}\right) - 1\right] \cdot \frac{1}{i}\frac{d}{dt} = \frac{2i\boldsymbol{C}_{\boldsymbol{G}}}{(1+i\overrightarrow{\boldsymbol{G}})(1-i\overleftarrow{\boldsymbol{G}})} \cdot \frac{1}{i}\frac{d}{dt}$$

$$= 1 \cdot \boldsymbol{W} - \left(\frac{1-i\overrightarrow{\boldsymbol{G}}}{1+i\overrightarrow{\boldsymbol{G}}}\right)\left(\frac{1+i\overleftarrow{\boldsymbol{G}}}{1-i\overleftarrow{\boldsymbol{G}}}\right) \cdot \boldsymbol{H}', \qquad (A \cdot 16)$$

that is,

$$2\boldsymbol{C}_{\boldsymbol{G}} \cdot \frac{d}{dt} = (1+i\overrightarrow{\boldsymbol{G}})(1-i\overleftarrow{\boldsymbol{G}}) \cdot \boldsymbol{W} - (1-i\overrightarrow{\boldsymbol{G}})(1+i\overleftarrow{\boldsymbol{G}}) \cdot \boldsymbol{H}', \qquad (A \cdot 17)$$

then we find directly the simplest result which is given in the text, eq. (46).

References

1) Y. Nambu, Prog. Theor. Phys. **4** (1949) 331, 399.
2) Hausdorf, Leipziger Berichte **58** (1906), 19.
3) R. P. Feynman, lecture note at C. I. T. (1950), Phys. Rev. **84** (1951), 108; and
   I. Fujiwara, Prog. Theor. Phys. **7** (1952), 433.
4) Y. Nambu, private communication in July 1949, and
   O. Klein, ZS. f. Physik, **80** (1934), 792.
5) L. L. Foldy and S. A. Wouthuysen, Phys Rev. **78** (1950), 29 ; and
   S. Tani, Prog. Theor. Phys. **6** (1951), 267.
6) L. L. Foldy, Phys. Rev. **84** (1952), 168.

On the Bose-Einstein Liquid Model for Liquid Helium, II
— *Properties under High Pressures* —

Ziro MIKURA

*The Research Institute for Iron, Steel and Other Metals,
Tohoku University, Sendai*

The excitation energy in the "modified Bose-Einstein liquid" model is assumed to be proportional both to the number density of He^4 particles and to one third power of the particle mass. In a previous paper, many properties of the He^3- He^4 liquid mixture were accounted for by using this model, assuming a particle mass independent of He^3 concentration. In the present paper, this model is shown to be successfully applied also to the case of pure liquid He^4 under high pressures. The values of parameters (the excitation energy and the particle mass), which are chosen so as to reproduce the correct lambda-temperature corresponding to the pressure concerned, are found to be of the correct magnitude to explain the experimental pressure dependences of second sound velocity and of the entropy on the lambda-curve. A noticeable fact is that the increase in the value of the particle mass due to pressure is almost exactly proportional to pressure.

§ 1.   Introduction

A " modified Bose-Einstein liquid " theory was developed and successfully applied to the problem of He^3-He^4 mixtures in a previous paper[1] (referred to hereinafter as **I**). The essential point of this theory is to take the excitation energy $\varDelta$ as proportional to the number density of He^4 particles. If the same idea is applied to the case of liquid He^4 under high pressures, one is led to assume the excitation energy proportional to the liquid density. In the case of the mixture, it was unnecessary to assume a mass of the He^4 particle different from that for pure He^4 liquid. It is likely, however, that the particle mass is dependent on pressure. On the other hand, it is very probable that the excitation energy depends on pressure not only through the number density of the particles but also through their individual mass M_4. Thus, it appears plausible to assume an excitation energy proportional both to the density of the liquid and to some power of the particle mass. Concerning the velocity of second sound under pressure, nearly the best fit of the theory to observation is obtained if this power is chosen as $1/3$. This specially appointed relation between $\varDelta$, ρ, and M_4 determines the parameters $\varDelta$ and M_4 uniquely under the condition that the calculated transition temperature should be equal to the observed one. A very remarkable fact is that the increase in the mass M_4 is almost exactly proportional to pressure. The normal fraction, the entropy, and the specific heat at high pressures are calculated by employing these parameters. The calculated increase of entropy

due to pressure on the lambda-curve is compared with the values computed by Keesoms[2] from their observed values of ρ and $d\rho/dT$ under pressure. Agreement between theory and experiment is very satisfactory.

§ 2. Velocity of second sound under high pressures

The phenomenon of the second-sound wave propagation under high pressures seems to be essentially similar to that under the saturated vapour pressure. The normal fraction x, the entropy S, the specific heat C_v, and the velocity of second sound v_{II} (at low frequencies) are given by

$$x=\left(\frac{2\pi M_4 kT}{h^2}\right)^{3/2}\nu\,\frac{V}{N_4}\exp\,(-\varDelta/T),\tag{1}$$

$$S=\frac{k}{m_4}\left(\frac{2\pi M_4 kT}{h^2}\right)^{3/2}\frac{V}{N_4}\left(\frac{\varDelta}{T}+\frac{5}{2}\right)\exp(-\varDelta/T).\tag{2}$$

$$C_v=\frac{k}{m_4}\left(\frac{2\pi M_4 kT}{h^2}\right)^{3/2}\frac{V}{N_4}\left(\frac{\varDelta^2}{T^2}+\frac{3\varDelta}{T}+\frac{15}{4}\right)\exp(-\varDelta/T),\tag{3}$$

$$v_{II}^2=\frac{(1-x)\,TS^2}{xC_v,}\tag{4}$$

where $\varDelta$ and $\nu=M_4/m_4$ (the mass factor) should be considered as dependent on pressure. m_4 is the mass of a He^4 atom and the other quantities have the usual meanings.

In the case of the He^3-He^4 mixture treated in I, $\varDelta$ was found to be proportional to the number density of He^4 particles while ν was independent of He^3 concentration. Hence it is plausible to assume $\varDelta$ proportional to the number density particles also in the case of liquid He^4 under high pressures, though it varies in this case with pressure. As ν seems to depend on pressure, $\varDelta$ is likely to depend on ν, not only through the number density of He^4 particles, but also through a factor M_4^τ ($\tau:a$ constant number to be determined). It is then

$$\varDelta=\varDelta_0\,\frac{N_4}{\nu V}\Big/\frac{N_4^0}{\nu_0 V_0}\cdot\left(\frac{M_4}{M_4^0}\right)^\tau,\tag{5}$$

where V is the volume of the liquid containing N_4 atoms of He^4 under a certain pressure and the quantities with the symbol (0) refer to some standard state, for which may be chosen, e.g., the normal lambda-point. Eq. (5) can be interpreted as a general law which applies to any case where the number density and/or the individual mass of He^4 particles differ from the standard values. In the special case of pure He^4 under pressure, it becomes

$$\varDelta=\varDelta_0\frac{\rho}{\rho_0}\left(\frac{\nu_0}{\nu}\right)^{1-\tau}.\tag{6}$$

On the other hand, putting $x=1$ and $T=T_\lambda$ in eq. (1) gives

$$1=\left(\frac{2\pi M_4 kT_\lambda}{h^2}\right)^{3/2}\nu\,\frac{V}{N_4}\exp(-\varDelta/T_\lambda),\qquad(7)$$

from which the lambda-temperature is found as a function of $\varDelta$ and ν. Inserting the observed value into T_λ in (7) and combining it with (6) give $\varDelta$ and ν as functions of pressure, provided that the value of γ is known. We determine γ so as to make the calculated velocities of second sound agree with those observed as well as possible at the experimentally used highest pressure (25 atm.). Our calculation leads to $\gamma=1/3$. Substituting it into (5) gives

$$\varDelta=\varDelta_0\frac{N_4/V}{N_4^0/V_0}\left(\frac{\nu_0}{\nu}\right)^{2/3}\qquad(8)$$

for the general case, and in particular, for pure liquid He^4 under pressure, it gives

$$\varDelta=\varDelta_0\frac{\rho}{\rho_0}\left(\frac{\nu_0}{\nu}\right)^{2/3}.\qquad(9)$$

Assuming the same values for ν_0 and $\varDelta_0$ as used in **I** ($8\cdot8$ and $8.609°K$ respectively), ν and $\varDelta$ are determined and the velocity of second sound is computed. The results are compared with Maurer and Herlin's[1] observation in Fig. 1. The agreement is not so good, even at $p=25$ atm. This implies that a better agreement cannot be expected so long as one uses the "modified B. E. liquid" model with temperature-independent parame-
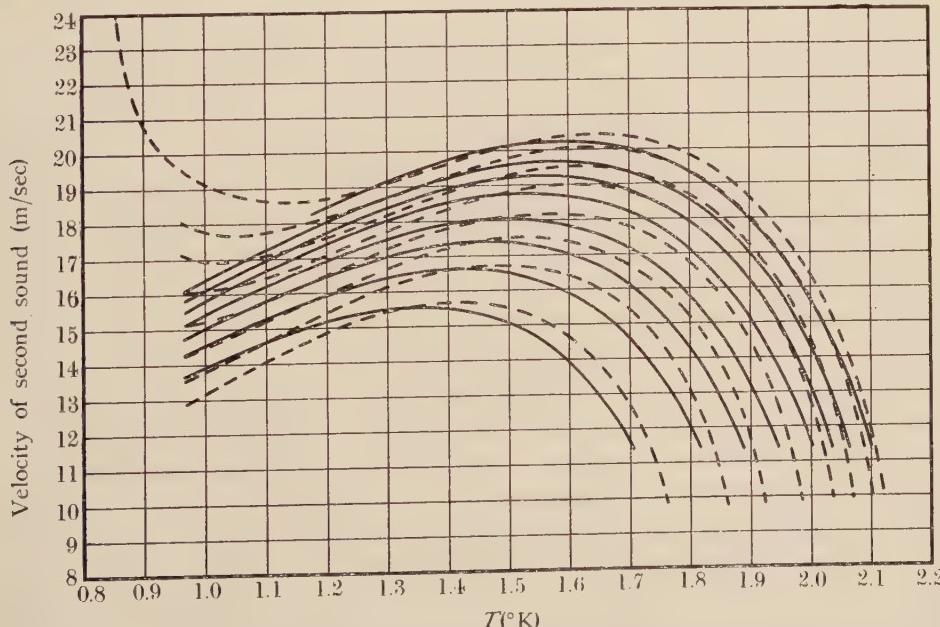
Fig. 1. The velocity of second sound as a function of temperature at various pressures. Solid curves are theoretical. Dashed curves indicate Maurer and Herlin's[1] smoothed experimental data.

ters, ν and J.   Nevertheless, it should be emphasized that the use for γ of a single value, 1/3, enables one to get nearly the best fit at any pressure.   The characteristic feature that theory gives higher values than experiment at low temperatures and lower ones at high temperatures is found not only at high pressures but also under the saturated vapour pressure.   It should be regarded as inherent to the present model with the temperature-independent parameters.   Thus eq. (9) gives nearly the best relation to determine ν and J as far as tne present model is concerned.   The numerical values of ν and J at various pressures are listed in Table I, together with those of some other quantities employed for evaluating them.   Those latter values are all taken or derived by interpolation from Keesom's monograph[3].

Table I. ν and J at various pressures, together with ρ and T_λ used for evaluating them.

p (atm)	0.05	2.7	4.7	7.1	10.8	14.5	18.8	25.0
ρ (g/cc)	0.1462	0.1505	0.1536	0.1572	0.1620	0.1661	0.1703	0.1757
T_λ(°K)	2.186	2.154	2.129	2.098	2.047	1.997	1.933	1.835
ν	8.800	9.192	9.493	9.863	10.428	10.975	11.650	12.714
J(°K)	8.609	8.608	8.599	8.579	8.519	8.442	8.318	8.096

§ 3.   Normal fraction, entropy and specific heat under high pressures

Using the proper values of ν and J (derived from (7) and (9)) corresponding to each pressure and applying eqs. (1), (2) and (3), we can calculate the normal fraction x, the entropy S, and the specific heat C_v under various pressures.   The relative increase due to pressure in these quantities are listed in Table II, together with the values of ν and J used.   Unfortunately, no direct measurements concerning these quanties have yet been made under high

Table II.   Relative increase due to pressure in x, S, and C_v (theoretical).

T(°K)	$p=5$ atm			$p=10$ atm			$p=15$ atm			$p=20$ atm			$p=25$ atm		
	x/x_0	S/S_0	C_v/C_v^0	x/x_0	S/S_0	C_v/C_v^0	x/x_0	S/S_0	C_v/C_v^0	x/x_0	S/S_0	C_v/C_v^0	x/x_0	S/S_0	C_v/C_v^0
1.0	...	...	...	...	...	...	1.86	1.45	1.43	2.49	1 80	1.73	3.49	2.30	2.18
1.2	1.17	1.08	1.08	1.44	1.22	1.21	1.80	1.41	1.39	2.36	1.70	1.65	3.20	2.12	2.01
1.4	1.17	1.08	1.08	1.42	1.21	1.20	1 76	1.38	1.36	2.27	1.64	1.58	3.01	2.00	1.89
1.6	1.17	1.08	1.08	1.41	1.20	1.19	1.73	1.36	1.34	2.20	1.59	1.54	2.88	1.91	1.82
1.8	1.17	1.08	1.08	1.41	1.19	1.19	1.71	1.35	1.32	2.15	1.56	1.51	2.78	1.85	1.76
2.0	1.17	1.08	1.08	1.40	1.19	1.18	...	...	...	...	...	...	...	...	...
ν	9.542			10.306			11.049			11.845			12.714		
J(°K)	8.597			8.533			8.429			8.280			8.096		

 *   Note added in proof: After this work was submitted, I became aware of the experimental work by W. H. Keesom and A. P. Keesom (Physica **2** (1935), 557) on the specific heat at high pressures.   The present theory has been found to be in good agreement with their results.   Details will appear in due course.

pressures.*   A rough estimation of entropy under pressure was made by Keesom and Keesom[2] from the observed values of ρ and $d\rho/dT$ under pressure.   The entropies on the lambda-curve are derived from their entropy diagram (Fig. 1 of reference[2]) and compared with the theoretical results, which are obtained by using the formula

$$S_\lambda = \frac{R\varDelta}{4\nu}\left(\frac{1}{T_\lambda} + \frac{5}{2\varDelta}\right). \tag{10}$$

This relation is easily derived from (1) and (2) by substituting $x=1$ and $T=T_\lambda$. Table III shows the values of $S-S_\lambda{}^0$ ($S_\lambda{}^0$ is referred to the normal lambda-point, $T=2.186°$K and $p=0.05$ atm.) calculated by making use of ν and $\varDelta$ in Table II, together with the corresponding values deduced from Keesoms' entropy diagram.   Agreement between the theoretical and the semi-empirical values is almost complete.

Table III. $S_\lambda - S_\lambda{}^0$ as a function of pressure.

p (atm.)		0.05	5	10	15	20	25
$S_\lambda - S_\lambda{}^0$	calc.	0.000	−0.023	−0.043	−0.061	−0.077	−0.093
(cal/g deg)	obs.	0.000	−0.023	−0.048	−0.068	−0.083	−0.095

The theoretical value of $S_\lambda{}^0$ is 0.364 cal/g deg, while the value derived by Kramers et al.[5] from their experimental data is 0.375 cal/g deg.

§ 4.   Discussion

The "modified B.E. liquid" model has been shown to be able to account for most observed properties of liquid He[4] under pressure if ν and $\varDelta$ are chosen so as to satisfy eq. (8) and to give the same lambda-temperture as observed experimentally.   The dependence of the excitation energy on the mean particle distance and the particle mass in the form given by (8) seems to support the suggestion the excitation energy originates from some kind of interactions between He[4] particles.   The values of ν in Table I are now plotted in Fig. 2 as a function of pressure.   It is remarkable that these values lie almost exactly on a single straight line, which can be represented by
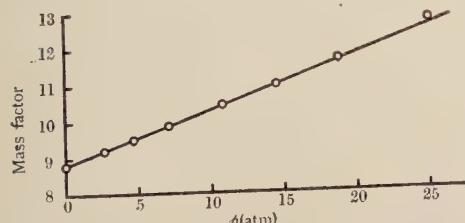
Fig. 2.   The mass factor, ν, as a function of pressure.   A straight line, $\nu=8.8+0.15$ (p −0.05), is also drawn.

$$\nu = 8.8 + 0.15\,(p - 0.05), \tag{11}$$

where p is the pressure in atm. It is desirable to establish any molecular theory which leads to the "modified B.E. liquid" model with ν and $\varDelta$ satisfying those two relations (8) and (11).   Most properties of liquid He[4] (except those related to the equation of state), both pure and diluted with He[3], under high pressures as well as under its saturated vapour pressure will be then well understood.

The possible existence of an energy gap at the bottom of the one-particle energy spectrum might be experimentally investigated if the microwave absorption experiment were able to be carried out in the neighbourhood of the corresponding wave-length. This wave-length lies in the neighbourhood of 1.7 mm (corresponding to the value, $8.6°K$, of $\varDelta_0$) for pure liquid He^4, and it seems difficult to carry out an absorption measurement using such a short radio-wave. In the case of He^3-He^4 mixtures, however, the circumstances are more favourable as the wave-length for the beginning of resonance absorption becomes longer. In a 50 percent solution of He^3 in He^4, for example, it turns out to be about 4 mm (corresponding to the value, $3.65°K$, of $\varDelta$)* and the measurement seems to be quite feasible.

Thanks are due to my colleagues, Messrs. A. Morita, N. Honda, C. Horie and Y. Fukuda, for their valuable advices.

References

1) Z. Mikura, Prog. Theor. Phys. **11**. (1954), 25.
2) W. H. Keesom and A. P. Keesom, Physica **1** (1933-4), 161.
3) W. H. Keesom, *Helium* Elsevier, Amsterdam, (1942).
4) R. D. Maurer and M. A. Herlin, Phys. Rev. **73** (1948), 729.
5) H. A. Kramers, J. D. Wasscher and C. J. Gorter, Physica **18** (1952), 329.

---

* See Paper **I**.

An Invariance Theorem for Cross Sections of Meson-Nucleon Scattering

Shigeo MINAMI

Department of Physics, Osaka University

We have certain points to bear in mind when we try to determine the values of phase shifts from experimental data for meson-nucleon scattering. Even if we could obtain an exact knowledge of the differential cross section for this process and, in analysing this, take into account high orbital angular momenta of meson, it will be found that two equally possible sets of solutions are in existence as phase shift. This ambiguity is of a fundamental character, has its origin in certain properties of the spherical harmonics and of the spin function and apparently remains unnoticed in the numerical analysis of Fermi and Yang.

§ 1. Introduction and summary

In order to study the meson-nucleon interaction, meson scattering by nucleon is one of the most important processes. In fact recent experimental results have given us considerably wealthy information. For the sake of interpretation of these results, the phase shift analysis has been adopted.[1)–5)] On the assumption that s- and p-waves only play the most important role, the phase shifts have been analysed by a number of authors. The values reported, however, are different from each other. This situation, of course, may mostly be due to the lack of sufficient data with respect to angular distributions. Moreover the higher the incident energy becomes, the more incorrect will be such an assumption of predominant s- and p-waves. It is the purpose of this paper to point out that, even if the angular distributions for this process be exactly found and a more elaborate method of analysis be applied, there remains an essential ambiguity, in fact there are always two equally possible sets of values for the phase shifts to account for one and the same angular distributions.

Let us classify the states of meson-nucleon system by the total angular momentum J and the total isotopic spin I. Each J-state consists of two different states which are characterized by parity. We represent these by the notations of R_J and S_J and the corresponding phase shifts by $\alpha_J(I)$ and $\beta_J(I)$ respectively.

Then our conclusion is: " the differential cross section $d\sigma/d\omega$ expressed in terms of phase shifts is invariant under the simultaneous substitution $\alpha_J(I) \rightleftarrows \beta_J(I)$ for all the J's and for all the I's."

§ 2. General considerations

When we consider a system of nucleon and pseudoscalar meson, the states of total

angular momentum $J=l+1/2$ are divided into two by the way J is composed, that is, nucleon spin $1/2+$meson orbital l or nucleon spin $1/2+$meson orbital $(l+1)$. They are of different parities and will be denoted by $\boldsymbol{R}_J$ and $\boldsymbol{S}_J$ respectively.

Let us lay the z-axis along the direction of the incident meson and represent the nucleon spin functions which have z-component $1/2$ and $(-1/2)$ by α and β respectively. If the initial spin state of nucleon is α, the z-component of the total angular momentum is $1/2$. (The conclusion runs quite the same in the case where the initial nucleon spin state is β.)

The J-component among scattered waves is written as follows:

$$e_J(I)\,(\boldsymbol{R}_J)_{1/2}+f_J(I)\,(\boldsymbol{S}_J)_{1/2}, \tag{1}*$$

where

$$e_J(I)=e^{2i\alpha_J(I)}-1, \quad f_J(I)=e^{2i\beta_J(I)}-1, \tag{2}$$

$$(\boldsymbol{R}_J)_{1/2}=2\sqrt{\pi}\,(2l+1)^{1/2}(\sqrt{(l+1)/(2l+1)})\left[\left(\frac{l+1}{2l+1}\right)^{1/2}\alpha Y_{l,0}+\left(\frac{l}{2l+1}\right)^{1/2}\beta Y_{l,1}\right],$$

$$(\boldsymbol{S}_J)_{1/2}=2\sqrt{\pi}\,(2l+3)^{1/2}(-\sqrt{(l+1)/(2l+3)})$$
$$\cdot\left[-\left(\frac{l+1}{2l+3}\right)^{1/2}\alpha Y_{l+1,0}+\left(\frac{l+2}{2l+3}\right)^{1/2}\beta Y_{l+1,1}\right]. \tag{3}$$

The expressions inside the square brackets are determined by the rule of composition of angular momenta, and the coefficient preceding the square brackets $\sqrt{(l+1)/(2l+1)}$ (or $-\sqrt{(l+1)/(2l+3)}$) is derived from the weight factor by which $(\boldsymbol{R}_J)_{1/2}$(or $(\boldsymbol{S}_J)_{1/2}$) is contained in the initial state $\alpha Y_{l,0}$ (or $\alpha Y_{l+1,0}$).
Our spherical harmonics are normalized on the unit sphere.

$$Y_{l,m}=(-1)^m\sqrt{\frac{(2l+1)}{2}\frac{(l-|m|)!}{(l+|m|)!}}\,P_l^m(\cos\theta)\,\frac{1}{\sqrt{2\pi}}\,e^{im\varphi}. \tag{4}**$$

Thus we can express the differential cross section in the center of mass system as follows:

$$\frac{d\sigma}{d\omega}=\frac{\lambdabar^2}{4}\left|\sum_I c(I)\sum_J\{e_J(I)\,(\boldsymbol{R}_J)_{1/2}+f_J(I)\,(\boldsymbol{S}_J)_{1/2}\}\right|^2. \tag{5}$$

$\sum_I$ indicates the sum over the isotopic spin state $1/2$ and $3/2$, and the coefficients $c(I)$ are given by the following table.

Table

Reaction \ $c(I)$	$\pi^++p\to\pi^++p$	$\pi^-+p\to\pi^-+p$	$\pi^-+p\to\pi^0+n$
$c(1/2)$	0	2/3	$-\sqrt{2}/3$
$c(3/2)$	1	1/3	$\sqrt{2}/3$

*) The suffix $1/2$ of $(\ )_{1/2}$ means the z-component of the total angular momentum.
**) Hereafter, without any mention we write $P_l^m(\cos\theta)$ as P_l^m.

1. The proof of $|(R_J)_{1/2}|^2 = |(S_J)_{1/2}|^2$

From (3)

$$|(S_J)_{1/2}|^2 = \{4\pi(l+1)/(2l+3)\}[(l+1)|Y_{l+1,0}|^2 + (l+2)|Y_{l+1,1}|^2]$$

$$= (l+1)\left[(l+1)|P_{l+1}|^2 + \frac{1}{l+1}|P_{l+1}^1|^2\right],$$

which is reduced by means of the recurrence formula

$$P_{l+1}^1 = \frac{(l+1)}{\sin\theta}[P_l - \cos\theta \cdot P_{l+1}] \tag{6}$$

to

$$|(S_J)_{1/2}|^2 = \frac{(l+1)^2}{\sin^2\theta}[|P_{l+1}|^2 + |P_l|^2 - 2P_l P_{l+1} \cdot \cos\theta]. \tag{7}$$

On the other hand, from (3)

$$|(R_J)_{1/2}|^2 = \{4\pi(l+1)/(2l+1)\}[(l+1)|Y_{l,0}|^2 + l|Y_{l,1}|^2]$$

$$= (l+1)\left[(l+1)|P_l|^2 + \frac{1}{l+1}|P_l^1|^2\right].$$

Also in this case, owing to another recurrence formula of the Legendre functions

$$P_l^1 = \frac{(l+1)}{\sin\theta}[\cos\theta \cdot P_l - P_{l+1}], \tag{8}$$

we get

$$|(R_J)_{1/2}|^2 = \frac{(l+1)^2}{\sin^2\theta}[|P_{l+1}|^2 + |P_l|^2 - 2P_l P_{l+1} \cdot \cos\theta]. \tag{9}$$

Thus, as we have asserted,

$$|(R_J)_{1/2}|^2 = |(S_J)_{1/2}|^2. \tag{10}$$

For example, when we consider the states up to $J = 3/2$,

$$|s_{1/2}|^2 = |p_{1/2}|^2 = 1,$$

$$|p_{3/2}|^2 = |d_{3/2}|^2 = 3\cos^2\theta + 1. \tag{11}*$$

If, in particular, we estimate the angular distribution and the total cross section for this process neglecting the contributions from any other scattering than resonance one, it is quite immaterial whether only $(R_J)_{1/2}$ resonates or only $(S_J)_{1/2}$ resonates.

---

*) The suffix 1/2 or 32/ in eq. (11) shows the total angular momentum J.

2. On the interference terms

Let us examine the cross terms in eq. (5). The states of total angular momentum $J' = l' + 1/2$ is divided into two by the same way as in the case of $J = l + 1/2$.* Corresponding to $(R_J)_{1/2}$ and $(S_J)_{1/2}$, we now define $(R'_{J'})_{1\,2}$ and $(S'_{J'})_{1\,2}$ respectively. Then $(R'_{J'})_{1/2}$ and $(S'_{J'})_{1/2}$ are obtained by substituting l' for l in eq. (3).

By means of the relations of (6) and (8), we are able to derive the following results

$$(R_J)^{*}_{1/2}(R'_{J'})_{1/2} = (R_J)_{1/2}(R'_{J'})^{*}_{1/2} = (S_J)^{*}_{1/2}(S'_{J'})_{1\,2} = (S_J)_{1/2}(S'_{J'})^{*}_{1/2}$$

$$= \frac{(l+1)\,(l'+1)}{\sin^2\theta}\left[P_{l+1}P_{l'+1} + P_l\,P_{l'} - (P_l\,P_{l'+1} + P_{l'}\,P_{l+1})\cos\theta\right], \qquad (12)$$

$$(R_J)^{*}_{1/2}(S'_{J'})_{1/2} = (R_J)_{1/2}(S'_{J'})^{*}_{1/2} = (S_J)^{*}_{1/2}(R'_{J'})_{1\,2} = (S_J)_{1/2}(R'_{J'})^{*}_{1/2}$$

$$= \frac{(l+1)\,(l'+1)}{\sin^2\theta}\left[P_l\,P_{l'+1} + P_{l'}\,P_{l+1} - (P_l\,P_{l'} + P_{l+1}\,P_{l'+1})\cos\theta\right] \qquad (13)$$

and

$$(R_J)^{*}_{1/2}(S_J)_{1/2} = (R_J)_{1/2}(S_J)^{*}_{1/2} = \frac{(l+1)^2}{\sin^2\theta}\left[2P_l\,P_{l+1} - (|P_{l+1}|^2 + |P_l|^2)\cos\theta\right].$$

$$(14)*$$

For example, when we consider the states up to $J = 3/2$,

$$s_{1/2}\cdot p_{3/2} = p_{1/2}\cdot d_{3/2} = 2\cos\theta,$$

$$s_{1/2}\cdot d_{3/2} = p_{1/2}\cdot p_{3/2} = 3\cos^2\theta - 1. \qquad (15)**$$

It is remarkable that the inner product of α and β vanishes on account of the property of orthogonality and as a result the contribution from the cross term $Y^{*}_{l,m}\,Y^{-}_{l',m'}$ is zero except in the case $m\quad m'$***. Therefore the value of any one cross term turns out to be real as implied in (12), (13) and (14).

3. The expression for the differential cross section

Employing the results obtained in § 2. 1 and § 2. 2, we are able to express the differential cross section as follows

---

*) Note that l in eqs. (7), (9), (12), (13) and (14) is defined by $(J-1/2)$ and does not necessarily mean the orbital angular momentum of meson.

  **) The meaning of the suffix is the same with that of (11).

***) In our case m (or m') is equal to zero or one. (c. f. (3)).

$$\frac{d\sigma}{d\omega} = \frac{\lambdabar^2}{4 \sin^2 \theta} \Big[\sum_{I,I''} c(I) c(I'') \sum_{J,I''} \{ e_{J}^{*}(I) c_{J''}(I'') + f_{J}^{*}(I) f_{J''}(I'') \}$$

$$\times (l+1)(l'+1) \{ P_{l+1} P_{l'+1} + P_{l} P_{l'} - (P_{l} P_{l'+1} + P_{l'} P_{l+1}) \cos \theta \}$$

$$+ \sum_{I,I''} c(I) c(I'') \sum_{J,J'} \{ e_{J}^{*}(I) f_{J''}(I'') + f_{J}^{*}(I) e_{J''}(I'') \}$$

$$\times (l+1)(l'+1) \{ P_{l} P_{l'+1} + P_{l'} P_{l+1} - (P_{l+1} P_{l'+1} + P_{l} P_{l'}) \cos \theta \} \Big], \quad (16)$$

where $J = l + 1/2$ and $J' = l' + 1/2$.

From the expression of (16), we see the following important circumstances. The differential cross section $d\sigma/d\omega$ is invariant under simultaneous substitution

$$e_{J}(I) \rightleftarrows f_{J}(I) \quad (17)$$

for all the J's and for all the I's.

If one set of possible values of phase shifts is found from the measurement of $d\sigma/d\omega$, the other set of values which is provided by the substitution (17) is also a solution. We have to bear this in mind when we analyse the experimental data.

§ 3. Discussions

The conclusion reached above is independent of whether the meson is pseudoscalar or scalar. Since each state is classified by total angular momentum, total isotopic spin and parity, the substitution (17) may be considered as the transformation with regard to parity. But the above conclusion does not mean that the $d\sigma/d\omega$ is invariant when pseudoscalar meson is substituted by scalar meson or vice versa. The reason is that the values of both $e_{J}(I)$ and $f_{J}(I)$ are expected to be different depending on the meson type and meson-nucleon interaction.

It would be very interesting to investigate why such a conclusion as in § 2 is brought about. Is it to be traced to some fundamental principle or is it a result of some particular condition ? In any case it may be noted that our process of summing over spin variables of the final nucleon gives rise to the symmetry we have found.

On the other hand, in lower energy region the following relation between phase shift and orbital angular momentum is known to be

$$k^{2l+1} \cot \delta_{l} = \text{const.} \quad (18)$$

On the basis of this relation, there may be some possibility to decide in favour of one or the other of the two sets of solutions by examining the energy dependence of the phase shifts, although the values of phase shifts hitherto reported do not necessarily satisfy this relation.

In conclusion, the author should like to express his thanks to Prof. K. Husimi, Prof. Z. Koba and Prof. S. Hayakawa for their valuable discussions.

References

1) H. L. Anderson, E. Fermi, R. Martin and D. E. Nagle, Phys. Rev. **91** (1953), 155.
2) D. Bodansky, A. M. Sachs and J. Steinberger, Phys. Rev. **90** (1953), 997.
3) J. P. Perry and C. E. Angell, Phys. Rev. **91** (1953), 1289.
4) W. B. Fowler, R. M. Lea, W. D. Shephard, R. P. Shutt, A. M. Thorndike and W. L. Whittemore, Phys. Rev. **92** (1953), 832.
5) S. W. Barnes, C. E. Angell, J. P. Perry, D. Miller, J. Ring and D. Nelson Phys. Rev. **92** (1953), 1327.

β–γ Angular Correlation of Tm170

Jun-ichi FUJITA, Masato MORITA and Masami YAMADA

Department of Physics, Faculty of Science, University of Tokyo

β–γ angular correlation of Tm170 is investigated taking into account the Coulomb field. Both the β-γ angular correlation and β-ray spectrum can be explained by the linear combination of the first forbidden nuclear matrix elements $\mathfrak{M}(\beta r)$, $\mathfrak{M}(\beta a)$ and $\mathfrak{M}(\beta\sigma \times r)$ in ST* type interaction of the Fermi theory of β-decay, and it is concluded that the changes of the spin and parity in the successive transitions of Tm170⟶Yb170*⟶Yb170 are $1(-) - 2(+) - 0(+)$.

Recently, R. Richmond and H. Rose[1] have investigated the β–decay of Tm170 experimentally. The decay scheme is shown in Fig. 1. Also H. Rose[2] measured the β–γ angular correlation of Tm170 for various energy values of electrons (Fig. 3). He pointed out that $\mathfrak{M}(\beta\sigma \times r)$ term** in T has a considerably good angular correlation in $z=0$ approximation[3], though he thought the fitness to be not complete. However, since the $\mathfrak{M}(\beta\sigma \times r)$ term, according to the formulation[4] of two of the present authors, is strongly affected by the Coulomb field, we shall in this paper attempt to explain the β–γ angular correlation by taking this effect into account, in the usual approximation of $(aZ)^2 \ll 1$***.
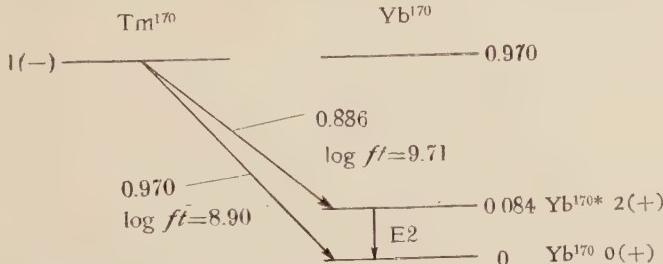
Fig. 1. Decay scheme of Tm170. Transition energies are given in Mev.

Both of the β–ray spectra for Tm170→Yb170 and Tm170→Yb170* (See, Fig. 1) have the allowed shapes[1][2] and log *ft*-values[5] 8.90 and 9.71 respectively. Therefore their orders of transition can not be forbidden higher than the first. The assumption of allowed transition of Tm170→Yb170* is excluded by the fact that the β–γ angular correlation exists in Tm170→

---

* The five interactions, Scalar, Vector, Tensor, Axialvector and Pseudoscalar are abbreviated as S, V, T, A and P, respectively. A linear combination of these interactions is written, for example, as ST, when the ratio of their coupling constants is not restricted to a special value.

** We use the same notation as in reference 4.

*** In the usual treatment of β–decay formulae, $aZ \ll 1$ means the omission of $(aZ)^n$ terms for $n \geq 2$.

$Yb^{170*} \rightarrow Yb^{170}$, while the allowed transition has no such correlation. It has been shown by R. J. Grant and R. Richmond,[6] and R. E. Bell and R. L. Graham[7] that the γ-ray in $Yb^{170*} \rightarrow Yb^{170}$ is an electric quadrupole radiation*, and so the spins and parities of Yb^{170*} and Yb^{170} are $2(+)$ and $0(+)$, respectively. Then Tm^{170} has one of the three possibilities $0(-)$, $1(-)$ and $2(-)$. If Tm^{170} has $0(-)$, the β-spectrum of $Tm^{170} \rightarrow Yb^{170*}$ has the a-type, contrary to the experiments, and the case of $2(-)$ of Tm^{170} is excluded by the same argument concerning the shape of the β-spectrum of $Tm^{170} \rightarrow Yb^{170}$. (In the former case the $\beta-\gamma$ angular correlation also contradicts the experiments, but in the latter case, the angular correlation may be accounted for by adjusting the magnitudes of nuclear matrix elements.) Thus the only remaining possibility for Tm^{170} is $1(-)$.

Recently it has become likely that both S and T exist in β-decay, and when the spin change is $1(-) \rightarrow 2(+)$, there are four reduced nuclear matrix elements $\mathfrak{M}(\beta r)$, $\mathfrak{M}(\beta \alpha)$, $\mathfrak{M}(\beta \sigma \times r)$ and $\mathfrak{M}(B_{ij}^{A})$. In the present treatment, we shall for simplicity use the first three only, because if their linear combination can explain the $\beta-\gamma$ angular correlation we shall also be able to explain it even though $\mathfrak{M}(B_{ij}^{A})$ term is included.

We introduce real parameters[4] x, y and z as

$$\frac{G_S \mathfrak{M}(\beta r)}{G_T \mathfrak{M}(\beta \sigma \times r)} = -ix, \quad \frac{\mathfrak{M}(\beta \alpha)}{\mathfrak{M}(\beta \sigma \times r)} = \frac{aZ}{2\rho} y \quad \text{and} \quad z = (x - y + 1)\frac{aZ}{2\rho}. \quad (1)$$

With these parameters, the β-ray angular distribution function $F_{11}^{0}(\theta)$ is greatly simplified and the large value $aZ/2\rho$ is apparently cancelled, and this saves largely the labor of calculation. From eqs. (S1), (T1) of YM and (ST1) of M, we have

$$F_{11}^{0}(\theta) = a_{11}^{(0)} - 2a_{11}^{(2)} P_2(\cos\theta)$$

$$= [z^2 + (2/3)\{K(1-x) + (p^2/W)(1+x)\}z$$

$$+ (K^2 + p^2)(1 + 2x^2)/6 + (2Kp^2/9W)(1-x^2)]$$

$$- (1-2x)[(2p^2/3W)z + (2Kp^2/9W)(1-x)$$

$$+ (p^2/6)(1+2x)]P_2(\cos\theta). \quad (2)$$

Substituting $a_{22}^{(2m)}$'s of the electric quadrupole radiation, eq. (24) of YM, into eq. (29) of YM, we get the $\beta-\gamma$ angular correlation function $\mathcal{W}(\theta)$ in the case of $1(-) - 2(+) - 0(+)$:

$$\mathcal{W}(\theta) = 1 - \{(3a_{11}^{(2)})/(4a_{11}^{(0)} + a_{11}^{(2)})\}\cos^2\theta. \quad (3)$$

The angular correlation coefficient $a(W)$ is defined by

$$\mathcal{W}(\theta) = 1 + a(W)\cos^2\theta. \quad (4)$$

For the maximum energy of the electron we use the value

---

* The life time of Yb^{170*} is about 10^{-9} sec,[8] and some reduction of angular correlation may occur owing to this rather long life of intermediate state. But, we neglect this effect in the following analysis.

$$W_0 = 2.74 \; mc^2.$$

We shall find such a domain of pairs of x and z, which give the good angular correlation and a good correction factor for the β-ray spectrum of $Tm^{170} \rightarrow Yb^{170}*$. This procedure is almost similar to the one which was used by the present authors in the case of Sb^{124} [9]. First, we put the theoretical value of $a(2.18)$ equal to the upper limit of its experimental value. This gives a restricting relation between x and z (Curve 1 in Fig. 2), and generally
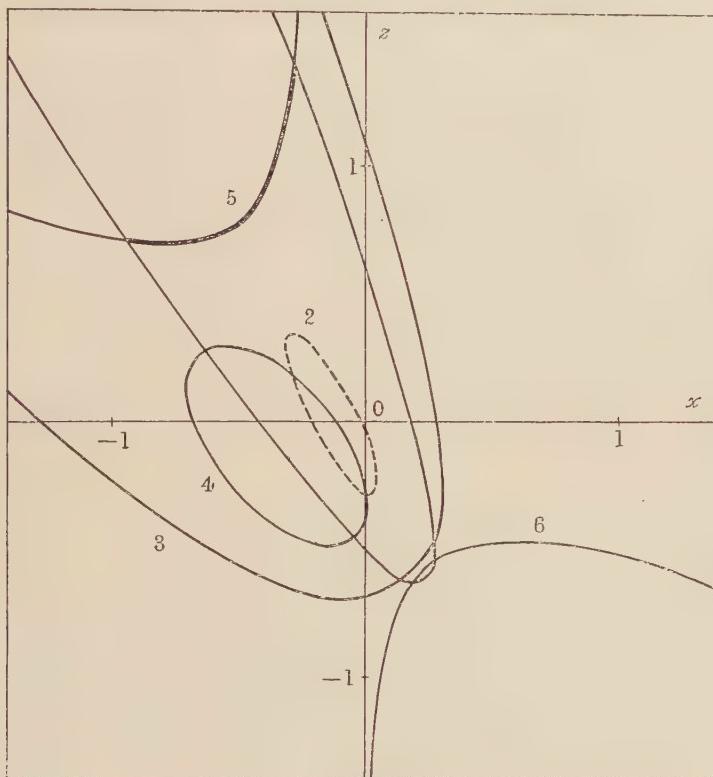
Fig. 2. Curves of the constant angular correlation coefficients for $1(-)-2(+)$ $-0(+)$, in the case of the linear combination of $\mathfrak{M}(\beta r)$, $\mathfrak{M}(\beta a)$ and $\mathfrak{M}(\beta \sigma \times r)$.

     1 : $a(2.18) = -0.24$,   2 : $a(2.18) = -0.32$,

     3 : $a(1.50) = -0.10$,   4 : $a(1.50) = -0.25$.

Curves 5 and 6, which give the same values of correction factors at $W = 1.50$ and 2.18 mc^2, are the two branches of a hyperbola.

they are the curves of the second degree. A Similar procedure with the lower limit of the experimental value of $a(2.18)$ yields another curve. This upper value is -0.34. But $a(2.18)$ $= -0.34$ has no real root, and so we give the curve $a(2.18) = -0.32$. (Curve 2 in Fig. 2). The domain between these two curves gives the pairs of x and z which have the good angular correlation at $W = 2.18$. The same method is repeated at another

electron energy $W=1.50$. The meet of these two domains will give the pairs of x and z which provide us with the good angular correlation at these two points of electron energy. We must examine by drawing the graph of the theoretical $a(W)$, whether this graph fits to the experimental data for every energy of the electron or not. Moreover, we impose a condition that makes the Kurie plot straight at three points of W ($W=1.50$ and $2.18\ mc^2$ and W_0), (Curves 5 and 6 in Fig. 2 which are the two branches of a hyperbola). Therefore, the points, which lie on the curves 5 or 6 and in the meet of the above described two domains simultaneously, will give the good angular correlation factor.

The angular correlation and correction factor are drawn in Fig. 3 and 4 for three points in Fig. 2, $x=0.2$, $z=-0.4$; $x=0.3$, $z=-0.5$ and $x=-0.4$, $z=1$. It seems that the neighboods of $x=0.2$, $z=-0.4$ and $x=0.3$, $z=-0.5$ give poor correction factors, while the meet to which $x=-0.4$, $z=1$ belong provides both the good angular correlation and the good correction factor.

The corrected ft-value for $\int\beta\sigma\times r$, $\mathrm{Tm}^{170}\to\mathrm{Yb}^{170}*$ is $ft=(2\pi^3/G_T^2)\log2/|\int\sigma\times r|^2\approx 1\times10^{10}$, ($x=-0.4$, $z=1$).

Since we have neglected $\mathfrak{M}(B_{ij}^A)$ term and other corrections, e.g. the finite nuclear size effect[10], it is not possible to deduce further definite conclusion. However, it is clear that $\mathfrak{M}(\beta\sigma\times r)$ term only can not explain the experiment (in this case $x=y=0$, $z=(\alpha Z/2\rho)\approx12.7$). When $\mathfrak{M}(B_{ij}^A)$ term is neglected, $\mathfrak{M}(\beta r)$ term, namely S, is necessary,
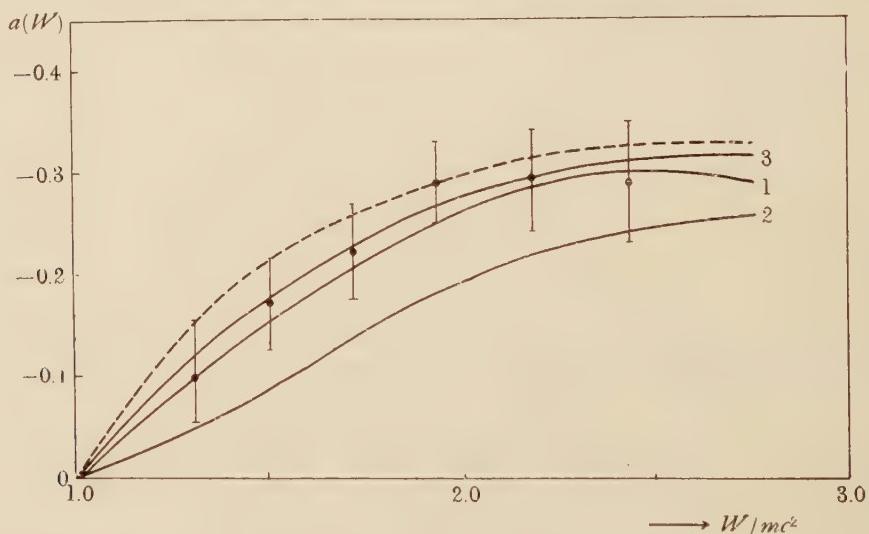
Fig. 3. Angular correlation coefficient for $1(-)-2(+)-0(+)$, in the case of the linear combination of $\mathfrak{M}(\beta r)$, $\mathfrak{M}(\beta\alpha)$ and $\mathfrak{M}(\beta\sigma\times r)$.

1: $x=0.2$, $z=-0.4$;   2: $x=0.3$, $z=-0.5$;

3; $x=-0.4$, $z=1$; Dotted curve: $\mathfrak{M}(\beta\sigma\times r)$ only in the approximation $Z=0$.

$\emptyset$: Experimental values of H. Rose.[2])
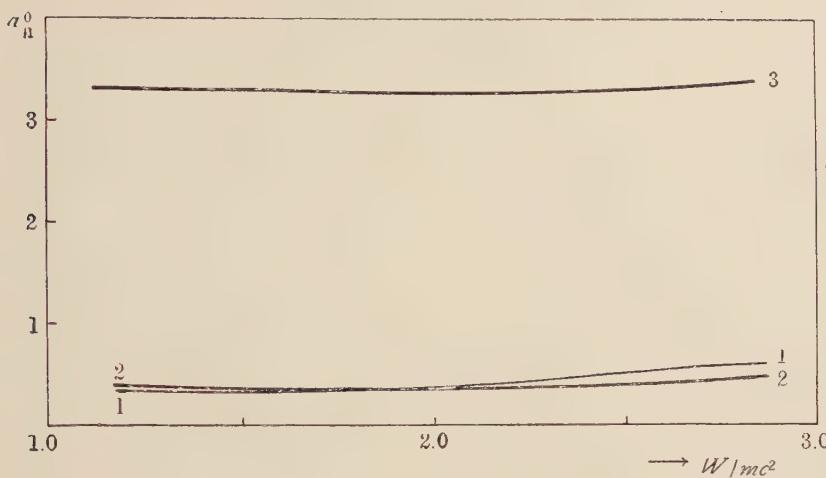
Fig. 4. Correction factors $(a^{(0)}_{13})$ in the case of the linear combination of $\mathfrak{M}(\beta r)$, $\mathfrak{M}(\beta a)$ and $\mathfrak{M}(\beta\sigma\times r)$.

         1: $x=0.2$, $z=-0.4$;   2: $x=0.3$, $z=-0.5$;

         3: $x=-0.4$, $z=1$.

but $\mathfrak{M}(B^3_{ij})$ term may change this conclusion. Lately it has been concluded that the ratio of the β-decay coupling constants G_S/G_T has the minus sign[11]. Then, combining it with the minus sign of the parameter x, we get

$$\frac{i\mathfrak{M}(\beta r)}{\mathfrak{M}(\beta\sigma\times r)} > 0.$$

This might supply some information about the structures of Tm^{170} and $Yb^{170}*$ together with their spins $1(-)$ and $2(+)$ respectively.

     The authors wish to express their sincere thanks to Professors T. Yamanouchi and S. Nakamura, Doctor M. Umezawa and Mr. H. Takebe for kind guidances and valuable discussions. They are also indebted to " Yukawa-Yomiuri Fellowship " for the financial aid to two of the authors (M. M. and M. Y.).

References

1) R. Richmond and H. Rose, Phil. Mag. **43** (1952), 367.

2) H. Rose, Phil. Mag. **43** (1952), 1146.

3) D. L. Falkoff and G. E. Uhlenbeck, Phys. Rev. **79** (1950), 334.

4) M. Yamada and M. Morita, Prog. Theor. Phys. **8** (1952), 431;
   M. Morita, Prog. Theor. Phys. **10** (1953), 363, called YM and M hereafter.

5) A. M. Feingold, Rev. Mod. Phys. **23** (1951), 10.

6) P. J. Grant and R. Richmond, Proc. Phys. Soc. **A62** (1949), 573.

7) R. E. Bell and R. L. Graham, Phys. Rev. **78** (1950), 490.

8) F. K. McGowan, Phys. Rev. **85** (1952), 151.

9) M. Morita and M. Yamada, Prog. Theor. Phys. **8** (1952), 449; **10** (1953), 111 and 641.

10) M. E. Rose and D. K. Holmes, Phys. Rev. **83** (1951), 190; Oak Ridge National Laboratory Report ORNL 1022 (1951).
    M. Yamada, Prog. Theor. Phys. **10** (1953), 245 and 241.

11) M. Morita, J. Fujita and M. Yamada, Prog. Theor. Phys. **10** (1953), 630. D. C. Peaslee, Phys. Rev. **91** (1953), 1447.

Letters to the Editor

On the Renormalization of Salpeter-Bethe Kernel of Meson-nucleon System in the State $T=1/2$

Tetz Yoshimura

Department of Physics,
Tokyo University of Education

February 3, 1954

Recently many authors have developed the non-perturbational methods for the problem of meson nucleon scattering.

Although those methods are not expressed in powers of the coupling constants, they are not free from divergences, so that it is necessary to renormalize them without using the expansion in powers of the coupling constants.

Some of their methods are not relativistically invariant, so it is impossible to remove divergences unambiguously. On the other hand, Fubini[1] has shown a subtraction method for the Salpeter-Bethe[2] kernel for the meson nucleon scattering problems in the state $T=1/2$, but his method is inconsistent because of the fact that two different subtractions are necessary for the nucleon propagation function and the vertex part, respectively. For the consistency of the renormalization, it is required that the solution of the integral equation can be expressed in terms of renormalized quantities multiplied by a constant factor as a whole. It seems to be a formal require-ment, but it is necessary to obtain a physically meaningful result from such a method.

Then, as a first step for consistent renormalization of those non-perturbational methods, we shall carry out the renormalization of Salpeter-Bethe formalism, which is most perturbation-like among them but of which relativistic invariance is guaranteed.

Let us start from such integral equation for the Salpeter-Bethe kernel with the "second order" inter-action functions, namely "ladder approximation", in which the inhomogeneous term is not $S_F \Delta_F$ but $S_F' \Delta_F'$, as demonstrated for the case of two fermion problem by Matthews and Abdus Salam;[3]

$$G(x\xi\,;\ x'\xi') = S_F'(x-x')\,\Delta_F'(\xi-\xi')$$

$$+\int dx_1 dx_2 d\xi_1 d\xi_2 S_{F}'(x-x_1)\,\Delta_{F}'(\xi-\xi_1)\cdot$$

$$\cdot\{I_1(x_1\xi_1\,;\ x_2\xi_2)+I_2(x_1\xi_1\,;\ x_2\xi_2)\}\,G(x_2\xi_2\,;\ x'\xi'). \tag{1}$$

The method, which will be developed below, can be easily extended to the case where higher order irreducible interactions are taken in I_2.

For convenience of further considerations we shall transform the equation into the momentum space;

$$G(p\,;\ kk') = S_{F}'(p-k)\,\Delta_{F}'(k)\,\delta(k-k')$$

$$+\int dk_1 S_{F}'(p-k)\,\Delta_{F}'(k)\{I_1(p)+I_2(p\,;\ kk_1)\}\cdot$$

$$\cdot G(p\,;\ k_1 k')\,, \tag{2}$$

where

$$I_1(p) = -g^2\tau\gamma_5 S_F(p,\ m_0)\,\tau\gamma_5, \tag{3}$$

$$I_2(p\,;\ kk') = -g^2\tau\Gamma_5(p-k,\ p-k-k')$$

$$\cdot S_{F}'(p-k-k')\,\tau\Gamma_5(p-k-k',\ p-k'), \tag{4}$$

$$S_{F}'(p) = S_F(p,\ m_0)\{1-\textstyle\sum^*(p)S_F(p,\ m_0)\}^{-1}, \tag{5}$$

$$\Delta_{F}'(k) = \Delta_F(k,\ \varkappa_0)\{1-\Pi^*(k)\,\Delta_F(k,\ \varkappa_0)\}^{-1}, \tag{6}$$

in the symmetrical $P_s\,(ps)$ meson theory and some numerical factors are discarded.

As it does not affect the consistency of the re-normalization to take the meson self-energy part Π^* in any approximation, we can take Π^* up to suitable order in g. (Similarly the renormalization of the so-called "square part" appearing in the higher order interaction functions can be carried out without affecting the consistency of renormalization.) But it destroys the consistency of the renormalization of G, to take $\sum^*$ and Γ_5 in arbitrary approximation. Then we shall consider, by what choice of $\sum^*$ and Γ_5 the renormalization of G can be made consistent. The conditions which must be satisfied by $\sum^*$ and Γ_5 for the consistency of the renormalization are;

$$\textstyle\sum^*(p) = -g^2\int dk\,\gamma_5 S_{F}'(p-k)\,\Delta_{F}'(k)\,\Gamma_5(p-k,\ p)$$

$$= -g^2\int dk\,\Gamma_5(p,\ p-k)\,S_F(p-k)\,\Delta_{F}'(k)\,\tau\gamma_5, \tag{7}$$

$$\Gamma_5(p,\ p-k) = \int dk'\gamma_5 S_{F}'(p-k')\,\Delta_{F}'(k')\,R(p\,;\ k'k)$$

$$+\gamma_5 \tag{8}$$

$$= \int dk'\,R(p-k\,;\ -k,\ k')\cdot$$

$$\cdot S_{F}'(p-k-k')\,\Delta_{F}'(k')\,\gamma_5 + \gamma_5,$$

where R is the resolvent of I_2 ;

$$R(p, kk') = I_2(p ; kk') + \int dk_1 I_2(p ; kk_1)$$
$$\cdot S_{F'}(p-k_1) \Delta_F(k_1) R(p ; kk'). \qquad (9)$$

These conditions should be satisfied by exact $\sum^*$, Γ_5 and I_2 automatically, and it can be easily shown that no overlap of diagrams occurs in the solutions of these equations if expanded in powers of g.

Following the usual method, we carry out the renormalization of the one body propagation functions, the vertex part and the coupling constant ;

$$S_{F'} = Z_2 S_{F_1'}, \qquad \Delta_{F'} = Z_3 \Delta_{F_1'},$$
$$\Gamma_5 = Z_1^{-1} \Gamma_{51}, \qquad g = Z_1 Z_2^{-1} Z_3^{-1/2} g_1, \qquad (10)$$

where Z's are infinite constants defined as following ;

$$Z_2 = (i\gamma p + m) S_{F'}(p) |_{i\gamma p + m = 0}, \qquad (11)$$

$$Z_3 = (k^2 + \varkappa^2) \Delta_F(k) |_{k^2 + \varkappa^2 = 0}, \qquad (12)$$

$$Z_1^{-1} = \gamma_5 \Gamma_5(p, p) |_{i\gamma p + m = 0}, \qquad (13)$$

where $\qquad \delta m = \sum^*(p) |_{i\gamma p + m = 0}, \qquad (14)$

$$\delta \varkappa (\delta \varkappa + 2\varkappa_0) = \Pi^*(k) |_{k^2 + \varkappa^2 = 0}. \qquad (15)$$

Now, Z's, δm, $\delta \varkappa$ are functions of g, m_0 and $\varkappa_0$, which can not be observed directly. Thus we must fit the values of these parameters by following conditions ;

$$[S_{F'}(p)]^{-1} |_{i\gamma p + m = 0} = 0, \qquad (16)$$

$$[\Delta_{F'}(k)]^{-1} |_{k^2 + \varkappa^2 = 0} = 0, \qquad (17)$$

$$g_1 = Z_1^{-1} Z_2 Z_3^{1/2} g, \qquad (18)$$

where m, $\varkappa$ and g_1 are observed values, and we must rewrite $S_{F_1'}$, $\Delta_{F_1'}$, Γ_5 and R in terms of m, $\varkappa$ and g_1.

Expressing the equation for R in terms of renormalized quantities ;

$$R(p ; kk') = (Z_2 Z_3)^{-1} I_2(p ; kk')$$
$$+ \int dk_1 I_{21}(p ; kk_1) S_{F_1'}(p-k_1) \Delta_{F'}(k_1)$$
$$\cdot R(p ; k_1 k'), \qquad (19)$$

where $\qquad I_{21} = -g_1^\circ \tau \Gamma_{51} S_{F_1'} \tau \Gamma_{51}, \qquad (20)$

then R should be renormalized as follows ;

$$R = (Z_2 Z_3)^{-1} R_1. \qquad (21)$$

The formal solutions of (1) is ;

$$G = S_{F'} \Delta_{F'} \delta + g^2 S_{F'} \Delta_{F'} \tau \Gamma_5 S_{F'} \tau \Gamma_5 S_{F'} \Delta_{F'}$$
$$+ S_{F'} \Delta_{F'} R S_F \Delta_{F'}, \qquad (22)$$

so G is expressed in terms of renormalized quantities and renormalization constants ;

$$G = (Z_2 Z_3) [S_{F_1'} \Delta_{F_1'} \delta$$
$$+ g_1^2 S_{F_1'} \Delta_{F_1'} \tau \Gamma_{51} S_{F_1'} \tau \Gamma_{51} S_{F_1'} \Delta_{F_1'}$$
$$+ S_{F_1'} \Delta_{F_1'} R_1 S_{F_1'} \Delta_{F_1'}]. \qquad (23)$$

Thus we can define a physically meaningful Green function G_1 in the following manner ;

$$G = Z_2 Z_3 G_1. \qquad (24)$$

Thus formal proof has been completed.

Because the uniqueness of m, $\varkappa_0$ and g giving the same m, $\varkappa$ and g_1 is not guaranteed, it is not clear that the remaining finite corrections is one valued functions of m, $\varkappa$ and g_1. Then there remain difficult problems to be solved, namely, how to make the theory free from divergences, how to solve the non-linear simultaneous integral equations, and by what conditions definite correction terms can be obtained from renormalization method without affecting the integral equations. Now it should be noticed that the renormalization is necessary even in a theory containing no diverging integrals, and subtraction step by step affects the integral equations.

This is only an example, how difficult it is to formulate a consistent renormalization procedure for non-perturbational methods.

More complicated circumstances would occur in the problems, e.g. γ-π production, in connection with the requirement of gauge invariance.

1) S. Fubini, Nuovo Cimento **X**, 6, (1953) 851.

2) E. E. Salpeter and H. A. Bethe, Phys. Rev. **84**, 6, (1951) 1232.

3) P. T. Matthews and Abdus Salam, unpublished (1953).

On Field Equations with Non-local Interaction

Chushiro Hayashi

Department of Physics, Naniwa University

February 11, 1954

In a previous paper[1] (to be referred to as I) it was shown that quantized fields with non-local interaction developed by Kristensen-Møller[2] and Bloch[3] could be fully described in the frame of Hamiltonian formalism which, on one hand, is suited to the bound-state problems and, on the other, enables us to construct a S-matrix consistent with the field equations. Moreover, it was found impossible in general to find a S-matrix directly according to the method of the above authors. Concerning these points further remarks will be made here.

First, the results of I will be briefly outlined with notes added to clarify our standpoint a little further. We start from q-number field equations

$$(\Box - M^2)\,\psi(x) = -g \int u(2)\,\psi(3)\,\varPhi(x, 2, 3)\,d(2\,3)\,,$$

$$(\Box - m^2)\,u(x) = -g \int \psi^*(1)\,\psi(3)\,\varPhi(1, x, 3)\,d(1\,3)\,,$$

$$\text{(1)}$$

and write them in the equivalent integral form of Yang-Feldman

$$\psi(x) = \psi(x, \sigma) - g \int \varDelta^\sigma(x-1)\,u(2)\,\psi(3)\,\varPhi(123)$$
$$\cdot d(123) + \varphi(x, \sigma)\,,$$

$$u(x) = u(x, \sigma) - g \int D^\sigma(x-2)\,\psi^*(1)\,\psi(3)\,\varPhi(1, 2, 3)$$
$$\cdot d(123) + \tau(x, \sigma)\,, \quad \text{(2)}$$

where $\psi(x, \sigma)$ and $\varphi(x, \sigma)$ both satisfy the free field equations, the same being for $u(x, \sigma)$ and $\tau(x, \sigma)$. The separation of $\psi(x, \sigma)$ from the additive term $\varphi(x, \sigma)$ is performed in such a way that $\varphi(x, \sigma)$ vanishes in the local case, and that $\psi(x, \sigma)$ satisfies free commutation relations for any fixed σ and, consequently, is connected with each other for fixed x and varying σ by unitary transformations which are generated by the interaction Hamiltonian. Such a procedure is possible for any field equations in principle without using series expansion if the form factor belongs to the normal class. In practice, however, calculations were carried out according to the series expansion in g, and it was found that $\varphi(x, +\infty)$

did not vanish in the fourth order when we chose $\varphi(x, -\infty) = v(x, -\infty) = 0$, that is, the construction of S-matrix according to the method of Kristensen-Møller was impossible. The overlook in the Bloch's proof concerning the existence of S-matrix lies in the fact that the conserving energy-momentum before and after collision, $G^\mu(in)$ and $G^\mu(out)$, do not both possess the same form as that of the free fields if they are constructed according to the *q-number* field equations. This will show that the variational principle for the Lagrangian is not so powerful as in the local theory. In our formalism, however, the out-field is defined by $\psi(x, +\infty)$ in (2) so that a unitary S-matrix consistent with the field equations (1) can be constructed. This interpretation of the out-field will be equivalent to a procedure proposed by Bloch* that with a knowledge of $\varphi(x, +\infty)$ we can always modify the right side of (1) by adding terms of order g^1 and higher in such a way that the new additive term vanishes for $\sigma = -\infty$ and $+\infty$.

Though there will be no difficulty in principle to obtain a S-matrix with our method in this way, it is desirable to find field equations which have a simple form and for which additive terms vanish for $\sigma = \pm 0$, being symmetric with respect to in- and out-fields. For this purpose the modification of the field equations is inevitable, and the simplest way for this will be to alter the order of operators appearing on the right of (1). In view of the correspondence to the local theory, one point is suggested that field equations should be symmetric with respect to charge, that is, to ψ and ψ^* or to real fields ψ_1 and ψ_2 defined by $\psi = (\psi_1 + i\psi_2)/\sqrt{2}$. Indeed, $\varphi(x, +\infty)$ shown by (42) in I has a form as expected to vanish by this symmetrization. Such a consideration will also be necessary in studying problems about the gauge invariance in the case of electro-magnetic interaction.

The above requirement is fulfilled if we symmetrize the right side of (1) completely as

$$u(2)\,\psi(3) \to \tfrac{1}{2}\{u(2)\,\psi(3) + \psi(3)\,u(2)\}\,,$$
$$\psi^*(1)\,\psi(3) \to \tfrac{1}{2}\{\psi^*(1)\,\psi(3) + \psi(3)\,\psi^*(1)\}\,,$$

$$\text{(3)}$$

and put a condition for the form factor

$$\varPhi(1, 2, 3) = \varPhi(3, 2, 1)\,. \quad \text{(4)}$$

In this case the interaction Hamiltonian and additive terms have been calculated up to the fourth order

---

* Comment given at the International Conference on Theoretical Physics held in Kyoto, Sept. 1953.

in the same way as in I. Using the Hermitian condition $\boldsymbol{\phi}^*(1, 2, 3) = \boldsymbol{\phi}(3, 2, 1)$ alone for the form factor, we have for $\varphi(x, \sigma)$ and $v(x, \sigma)$ which are chosen so as to vanish for $\sigma = -\infty$

$$\varphi(x, +\infty) = \frac{g^1}{8} \int \varDelta(x-1) f(1, 3''') d(1\ 3'''), \quad (5)$$

$$
\begin{aligned}
f(x, x''') = \int \varDelta(1'-3'') &\{ \varDelta^{\text{ret}}(3-1'') \\
\times D^{\text{ret}}(2''-2'') \varDelta^{\text{ret}}&(1'''-3') D^{\text{ret}}(2'-2) \\
- \varDelta^{\text{adv}}(3-1'') D^{\text{adv}}&(2''-2'') \varDelta^{\text{adv}}(1'''-3') \\
\times D^{\text{adv}}(2'-2) \} &\boldsymbol{\phi}(x, 2, 3) \boldsymbol{\phi}(1', 2', 3') \\
\times \boldsymbol{\phi}(1'', 2'', 3'') &\boldsymbol{\phi}(1''', 2''', x''') \\
\times d(23\ 1'\ 2'\ 3'\ 1''\ 2''\ &3''\ 1'''\ 2'''),
\end{aligned}
\quad (6)
$$

$$v(x, +\infty) = \frac{g^1}{8} \int D(x-2) h(2, 2''') u(2''') d(2\ 2'''), \quad (7)$$

$$
\begin{aligned}
h(x, x''') = \int D(2'-2'') &\{ \varDelta^{\text{ret}}(3-1'') \\
\times \varDelta^{\text{ret}}(3''-1''')) \varDelta^{\text{ret}}&(3'''-1') \varDelta^{\text{ret}}(3'-1) \\
- \varDelta^{\text{adv}}(3-1'') \varDelta^{\text{adv}}&(3''-1''') \varDelta^{\text{adv}}(3'''-1') \\
\times \varDelta^{\text{adv}}(3'-1) \} &\boldsymbol{\phi}(1, x, 3) \boldsymbol{\phi}(1', 2', 3') \\
\times \boldsymbol{\phi}(1'', 2'', 3'') &\boldsymbol{\phi}(1''', x''', 3''') \\
\times d(131'2'3'1''2''&3''1'''3''').
\end{aligned}
\quad (8)
$$

It is easily shown that (6) and (8) both vanish for any x and x''' if we require for the form factor further the invariance for reflection, $\boldsymbol{\phi}(1, 2, 3) = \boldsymbol{\phi}(-1, -2, -3)$, the reality condition (4), and the invariance for displacements as a whole. In (6) and (8) x and x''' are expressed by $X = (x+x''')/2$ and $r = x - x'''$, all the integration variables are displaced by X, their signs are reversed, and their names are interchanged (dash with two-dashes and no-dash with three-dashes in (6) and (8), and 1 with 3 in (6)), then expressions having an opposite sign are obtained. More exactly, (6) and (8) vanish when the form factor has an extra damping factor such as $\exp\{-\varepsilon|\ell_1 + \ell_2 + \ell_3|\}$ which makes the equations (2) well-defined, if we note that the limits are taken in the order: $\lim_{\varepsilon \to 0} \lim_{\sigma \to +\infty} \varphi(x, \sigma)$. Problems concerning the higher order terms are now under investigation. Finally, it is remarked that the S-matrix is not altered up to the second order by the above symmetrization.

1) C. Hayashi, Prog. Theor. Phys. **10** (1953), 533.
2) P. Kristensen and C. Møller, Det. Kong. Danske Vid. Sel. Mat.-fys. **27** (1952), No. 7.
3) C. Bloch, ibid. **27** (1952), No. 8.

On a Covariant Generalization of Tamm-Dancoff Approximation for Pion-nucleon Scattering

Kiyomi Itabashi

Physical Institute, Faculty of Science, Tohoku University

Among the various approximation methods so far proposed by many authors, the Tamm-Dancoff method[1] seems to be interesting. For, if one stands on the basis of the assumption that the pion-nucleon coupling is still not so large, it is natural to consider the average number of mesons around a nucleon to be rather small, and to ignore the effects of the configurations with many mesons. But, as is well-known, the T.D. method has been formulated in a non-covariant form, so that the consistent subtraction of divergences from the theory has been very difficult. So it is desirable to get a covariant generalization of the T.D. method, and several attempts for this have been already proposed.[2]

Since the concept of the meson numbers plays an important role in the original T.D. method, the straightforward generalization of this method would necessarily contain the space-like surface σ in the integral equations for the representatives of the state vector, as in the formalism by Cini[2]. However, such a formalism would give rise to a new type of divergences due to the sharpness of integrations.[3] Accordingly, to avoid this new difficulty, it is necessary to give up the concept of "number", and to substitute the more wide one for that. In this respect, the method used by Feynman[4] in his derivation of a covariant theory from the old quantum electrodynamics would be very instructive. Thus, we pay an attention to the "nucleon line" and trace the "development along the nucleon line" rather than the "development in time" or the "time-like development of events".

For simplicity, we will consider the pion-nucleon scattering. Neglecting all the diagrams that contain the nucleon loops (we have then only one connected nucleon line), and considering the emissions and absorptions of mesons along the nucleon line, we will get the simultaneous integral equations for the Green's functions K's with various meson numbers:

$$K(f_1 ; f_0, k_0^{\varepsilon})$$
$$= g S_F(f_1) \gamma_5 \tau_\alpha \int dk K(f_1 - k, k^\alpha ; f_0, k_0^{\varepsilon}),$$

$$K(f_1, k_1^\alpha ; f_0, k_0)$$
$$= \delta(f_1 - f_0) \delta(k_1^\alpha - k_0^{\varepsilon}) S_F(f_0) \Delta_F(k_0) +$$
$$+ g S_F(f_1) \Delta_F(k_1) \gamma_5 \tau_\alpha K(f_1 + k_1 ; f_0, k_0^{\varepsilon}) +$$
$$+ g S_F(f_1) \gamma_5 \tau_\beta \int dk K(f_1 - k, k_1^\alpha, k^\beta ; f_0, k_0^{\varepsilon}),$$

$$K(f_1, k_1^\alpha, k_2^\beta ; f_0, k_0^{\varepsilon})$$
$$= g S_F(f_1) \Delta_F(k_1) \gamma_5 \tau_\alpha K(f_1 + k_1, k_2^\beta ; f_0, k_0^{\varepsilon}) +$$
$$+ g S_F(f_1) \Delta_F(k_2) \gamma_5 \tau_\beta K(f_1 + k_2, k_1^\alpha ; f_0, k_0^{\varepsilon}) +$$
$$+ g S_F(f_1) \gamma_5 \tau_\tau \int dk K(f_1 - k, k_1^\alpha, k_2^\beta, k^\tau ; f_0, k_0^{\varepsilon}),$$

$$K(f_1, k_1^\alpha, k_2^\beta, k_3^\tau ; f_0, k_0^{\varepsilon})$$
$$= g S_F(f_1) \Delta_F(k_1) \gamma_5 \tau_\alpha$$
$$\times K(f_1 + k_1, k_2^\beta, k_3^\tau ; f_0, k_0^{\varepsilon}) +$$
$$+ g S_F(f_1) \Delta_F(k_2) \gamma_5 \tau_\beta$$
$$\times K(f_1 + k_2, k_1^\alpha, k_3^\tau ; f_0, k_0^{\varepsilon}) +$$
$$+ g S_F(f_1) \Delta_F(k_3) \gamma_5 \tau_\tau$$
$$\times K(f_1 + k_3, k_1^\beta, k_2^\beta ; f_0, k_0^{\varepsilon}) +$$
$$+ g S_F(f) \gamma_5 \tau_\delta$$
$$\int dk K(f_1 - k, k_1^\alpha, k_2^\beta, k_3^\tau, k^\delta ; f_0, k_0^{\varepsilon}),$$
$$\cdots\cdots \quad [\delta(k_1^\alpha - k_0^{\varepsilon}) \equiv \delta(k_1 - k_0) \delta_{\alpha\varepsilon}] \quad (1)$$

where $K(f_1, k_1^\alpha ; f_0, k_0^{\varepsilon})$, for example, is the Green's function for the scattering process in which the "initial" and "final" 4-momenta of the nucleon and meson are f_0, k_0 and f_1, k_1, respectively. The super-suffices of k's indicate the charge states of mesons. Here and in the following, the words "number", "initial" and "final" have the different meanings with the ordinary ones.

The similar equations have already been derived by Watanabe[5] from the definition of the Green's functions of Gell-Mann and Low.[6] His equations are more general than eqs. (1) here in that they include the closed nucleon loops and are applicable also to the many-nucleon problems. However, when the nucleon loops are included, the equations become non-linear[7] and very difficult to treat with, unless the adequate approximations for the meson propagators etc. are introduced. Moreover, it should be noted that the "initial" and "final" mesons are distinguished in our formalism, so that the precise form of our equations are different with those of Watanabe.

Because of the reason above mentioned, we would

neglect all the nucleon loops and take the equations (1) as the "generalized T. D. equations" for the pion-nucleon system. Thus, the "generalized T. D. approximation" is to treat the equations (1) omitting all the K's with the meson numbers larger than a certain number. For example, in the third order approximation where the meson "numbers" are restricted within 3, the quantities $K(f_1, k_1^\alpha, k_2^\beta, \cdots k_n^\delta ; f_0, k_0^{\varepsilon})$ are omitted for $n \geq 4$.

1) I. Tamm, Jour. Phys (U.S.S.R.) **9**, (1945) 449. S. M. Dancoff, Phys. Rev. **78**, (1950) 382.
2) M. Cini, Nuovo cimento **10**, (1952) 562 and 614. S. Fubini, Nuovo cimento **10**, (1953) 851.
3) H. Lehmann, Z. Naturforschung **89**, (1953) 579.
4) R. P. Feynman, Phys. Rev. **76**, (1949) 749 and 769.
5) I. Watanabe, Prog. Theor. Phys. **10**, (1953) 371.
6) M. Gell-Mann and F. Low, Phys. Rev. **84**, (1951) 350.
7) Cf. M. Neuman, Phys. Rev. **92**, (1953) 1021.

Renormalization in Generalized Tamm-Dancoff Approximation for Pion-nucleon Scattering

Kiyomi Itabashi

Physical Institute, Faculty of Science, Tohoku University

In the preceding note, we have get a covariant generalization of the Tamm-Dancoff method. The next problem is how to subtract the divergences from that generalized theory. Fubini (loc. cit.) has given the answer to this problem in the second order approximation. However, since the starting equation in this order of approximation is essentially a single integral equation and the procedure of Fubini rests on this specially simple circumstance, it is not applicable directly to the higher order approximations. The aim of this note is to show how to extend the Fubini's procedure so as to become applicable also to any higher order approximations.

For illustration, let us consider the third order approximation. Then, our equations reduce to 4 simultaneous equations as mentioned at the end of the preceding note. Substituting the last equation of them into the equation for $K(\hat{p}_1, k_1^\alpha, k_2^\beta; \hat{p}_0, k_0^\varepsilon)$ and performing the subtraction of the self-energy type divergence which appears from the above substitution, they become as follows:

$$K(\hat{p}_1; \hat{p}_0, k_0^\varepsilon)$$
$$= g S_F(\hat{p}_1) \gamma_5 \tau_\alpha \int dk K(\hat{p}_1 - k, k^\alpha; \hat{p}_0, k_0^\varepsilon),$$

$$K(\hat{p}_1, k_1^\alpha; \hat{p}_0, k_0^\varepsilon)$$
$$= \delta(\hat{p}_1 - \hat{p}_0) \delta(k_1^\alpha - k_0^\varepsilon) . S_F(\hat{p}_0) \Delta_F(k_0) +$$
$$+ g S_F(\hat{p}_1) \Delta_F(k_1) \gamma_5 \tau_\alpha K(\hat{p}_1 + k_1; \hat{p}_0, k_0^\varepsilon) +$$
$$+ g S_F(\hat{p}_1) \gamma_5 \tau_\beta \int dk K(\hat{p}_1 - k, k_1^\alpha, k^\beta; \hat{p}_0, k_0^\varepsilon),$$

$$K(\hat{p}_1, k_1^\alpha, k_2^\beta; \hat{p}_0, k_0^\varepsilon)$$
$$= g S_F{}'(\hat{p}_1) \Delta_F(k_1) \gamma_5 \tau_\alpha K(\hat{p}_1 + k_1, k_2^\beta; \hat{p}_0, k_0^\varepsilon) +$$
$$+ g S_F{}'(\hat{p}_1) \Delta_F(k_2) \gamma_5 \tau_\beta K(\hat{p}_1 + k_2, k_1^\alpha; \hat{p}_0, k_0^\varepsilon) +$$
$$+ g^3 S_F{}'(\hat{p}_1) \gamma_5 \tau_\tau \int dk S_F(\hat{p}_1 - k) \Delta_F(k_1) \gamma_5 \tau_\alpha$$
$$\times K(\hat{p}_1 + k_1 - k, k_2^\beta, k^\tau; \hat{p}_0, k_0^\varepsilon) +$$
$$+ g^3 S_F{}'(\hat{p}_1) \gamma_5 \tau_\tau \int dk S_F(\hat{p}_1 - k) \Delta_F(k_2) \gamma_5 \tau^\beta$$
$$\times K(\hat{p}_1 + k_2 - k, k_1^\alpha, k^\tau; \hat{p}_0, k_0^\varepsilon),$$

where the definition of $S_F{}'$ is same as that of Fubini (loc. cit.).

To solve these equations formally, we consider the scattering in which the " numbers " of meson vary as $2 \to 3 \to 2$, $3 \to 2$ $\cdots \to 2$.

$$R_2(\hat{p}_1, k_1^\alpha, k_2^\beta; \hat{p}_0, k_3^\tau, k_4^\delta)$$
$$= \delta(\hat{p}_1 - \hat{p}_0) S_F{}'(\hat{p}_0) \Delta_F(k_3) \Delta_F(k_4) [\delta(k_1^\alpha - k_3^\tau)$$
$$\times \delta(k_2^\beta - k_4^\delta) + \delta(k_1^\alpha - k_4^\delta) \delta(k_2^\beta - k_3^\tau)] +$$
$$+ g^3 S_F{}'(\hat{p}_1) \gamma_5 \tau_\varepsilon \int dk S_F(\hat{p}_1 - k) \Delta_F(k_1) \gamma_5 \tau_\alpha$$
$$\times R_2(\hat{p}_1 + k_1 - k, k_2^\beta, k^\varepsilon; \hat{p}_0, k_3^\tau, k_4^\delta) +$$
$$+ g^3 S_F{}'(\hat{p}_1) \gamma_5 \tau_\varepsilon \int dk S_F(\hat{p}_1 - k) \Delta_F(k_2) \gamma_5 \tau_\beta$$
$$\times R_2(\hat{p}_1 + k_2 - k, k_1^\alpha, k^\varepsilon; \hat{p}_0, k_3^\tau, k_4^\delta). \qquad (2)$$

Using this R_2, we define the kernel G_2 as follows:

$$G_2(\hat{p}_1, k_1^\alpha; \hat{p}_0, k_3^\tau)$$
$$= g^2 S_F(\hat{p}_1) \gamma_5 \tau_\beta \int d\hat{p}_2 \int dk_2 \int d\hat{p}_3 \int dk_4 \delta(\hat{p}_2 + k_2 - \hat{p}_1) \times$$
$$\times \delta(\hat{p}_3 + k_4 - \hat{p}_0) R_2(\hat{p}_2, k_1^\alpha, k_2^\beta; \hat{p}_3, k_3^\tau, k_4^\delta)$$
$$\gamma_5 \tau_\delta [\Delta_F(k_3)]^{-1}. \qquad (3)$$

Finally, we solve the following equation (4), which has G_2 as its kernel and represents the scattering process $1 \to (2 \to 3 \to 2 \to \cdots \to 2) \to 1 \to \cdots \to 1$.

$$R_1(\hat{p}_1, k_1^\alpha; \hat{p}_0, k_3^\tau) = G_2(\hat{p}_1, k_1^\alpha; \hat{p}_0, k_3^\tau) +$$
$$+ \int d\hat{p}_2 \int dk G_2(\hat{p}_1, k_1^\alpha; \hat{p}_2, k^\beta) R_1(\hat{p}_2, k^\beta; \hat{p}_0, k_3^\tau).$$
$$(4)$$

Then, it is easy to construct the formal solution of (1) in terms of R_1 and R_2, i.e.

$$K(\hat{p}_1; \hat{p}_0, k_0^\varepsilon) = [A(\hat{p}_1) - g^2 S_F(\hat{p}_1) \gamma_5 \tau_\alpha$$
$$\times \int dk \int dk_1 \int dp R(\hat{p}_1 - k, k^\alpha; \hat{p}, k_1^\beta)$$
$$\times \Delta_F(k_1) S_F(\hat{p}) \gamma_5 \tau_\beta]^{-1} \times$$
$$\times g S_F(\hat{p}_1) . [\gamma_5 \tau_\varepsilon S_F(\hat{p}_0) \Delta_F(k_0) \delta(\hat{p}_1 - k_0 - \hat{p}_0)$$
$$+ \gamma_5 \tau_\alpha \int dk R_1(\hat{p}_1 - k, k^\alpha; \hat{p}_0, k_0^\varepsilon) \Delta_F(k_0) . S_F(\hat{p}_0)],$$

$$K(\hat{p}_1, k_1^\alpha; \hat{p}_0, k_0^\varepsilon)$$
$$= \delta(\hat{p}_1 - \hat{p}_0) \delta(k_1^\alpha - k_0^\varepsilon) \Delta_F(k_0) . S_F(\hat{p}_0) +$$
$$+ R_1(\hat{p}_1, k_1^\alpha; \hat{p}_0, k_0^\varepsilon) \Delta_F(k_0) S_F(\hat{p}_0) +$$
$$+ g \Delta_F(k_1) S_F(\hat{p}_1) \gamma_5 \tau_\alpha K(\hat{p}_1 + k_1; \hat{p}_0, k_0^\varepsilon) +$$
$$+ \int d\hat{p}_2 \Lambda^{(1)}(\hat{p}_1, k_1^\alpha; \hat{p}_2) K(\hat{p}_2; \hat{p}_0, k_0^\varepsilon),$$

$$K(\hat{p}_1, k_1^\alpha, k_2^\beta; \hat{p}_0, k_0^\varepsilon)$$
$$= g \int d\hat{p}_2 \int d\hat{p}_3 \int dk_3 \int dk_4 \times$$
$$\times R_2(\hat{p}_1, k_1^\alpha, k_2^\beta; \hat{p}_2, k_3^\tau, k_4^\delta) \times$$
$$\times \gamma_5 \tau_\delta [\Delta_F(k_3)]^{-1} \delta(\hat{p}_2 + k_4 - \hat{p}_3) \times$$
$$\times K(\hat{p}_3, k_3^\tau; \hat{p}_0, k_0^\varepsilon), \qquad (5)$$

where

$$\Lambda^{(1)}(\hat{p}_1, k_1^\alpha; \hat{p}_2) = g \int d\hat{p} \int dk R_1(\hat{p}_1, k_1^\alpha; \hat{p}, k^\beta)$$
$$\times \Delta_F(k) S_F(\hat{p}) \gamma_5 \tau_\beta \delta(\hat{p} + k - \hat{p}_2)$$

$$A(\hat{p}_1) = 1 - g^2 S_F{}^*(\hat{p}_1) \gamma_5 \tau_\tau \int dk S_F(\hat{p}_1 - k),$$
$$\times \Delta_F(k) \gamma_5 \tau_\tau. \qquad (6)$$

The direct substitution of (5) into (1) shows that (5) is actually the solution. The physical meaning of the above procedure would become very clear, if we analyse the process in question by means of schematical method.

Thus, we have succeeded in finding the formal solution of our simultaneous integral equations by treating only the single integral equations. The renormalization procedure after this stage is essentially the same as that of Fubini. Of course, it is true

that the kernel G_2 includes the various types of divergence as well as the convergent parts, and the renormalization of such a function needs some careful treatments. But this is also not difficult if one makes use of the nature of K_2-equation. The above method would be also applicable to any higher order

approximations. The details of this and the preceding note will be reported elsewhere.

Concluding this note, the author wishes to thank to Professors K. Nakabayasi and I. Sato for their helpful guidance and kind encouragement.

ERRATA

Fourth Order Phase Shifts for Meson-proton Scattering in the Symmetrical Ps(Ps) Theory

K. Nakabayasi, K. Hasegawa and I. Yamamura

The following table should be inserted.

Table 1. Matrix elements $(iR \times 10^3)$ in C.M. system

energy (Mev)	2-nd order		4-th order					
	R_I	R_{II}	R_{I_a}	R_{I_b}	R_{I_c}	R_{II_a}	R_{II_b}	R_{II_c}
0	8.6	7.4	-0.32	-0.87	-0.87	-0.97	-1.7_5	-1.9
60	8.7	7.2	-0.29	-0.83	-0.83	$-0.81_5+i0.35$	$-1.7+i0.4$	$-1.9+i0.5$
110	8.8$_5$	7.1	-0.26	-0.79	-0.80	$-0.75+i0.45_5$	$-1.7+i0.6$	$-2.0+i0.7$
135	8.9	7.1	-0.25	-0.77	-0.79	$-0.72+i0.49_5$	$-1.7+i0.6$	$-2.0+i0.8$
180	9.1	7.0	-0.24	-0.74	-0.76	$-0.69+i0.56$	$-1.7+i0.7$	$-2.2+i1.0$

Note on Meson-nucleon Interaction

S. Minami

p. 123, Left column, 4 line   *for* iostpic,   *read* isotopic

„ , Right „ , 1 line   *for* Clebsh,   *read* Clebsch

p. 124, Left „ , expression in case (a)

$$\textit{for} \quad 3|Y_{1,1}|^2 + (1/3)2|Y_{1,1}|^2 + |Y_{1,-1}|^2],$$

$$\textit{read} \quad 3|Y_{1,1}|^2 + (1/3)[2|Y_{1,1}|^2 + |Y_{1,-1}|^2]$$

„ , „ , 15 line   *for* $(2+\sin^2\theta)$,   *read* $(2+3\sin^2\theta)$

„ , Right „ , 9 line   *for* d-meson,   *read* d-wave meson.

Anomalous Magnetic Moment of Nucleon and Nucleon Isobar

Akira KANAZAWA and Masao SUGAWARA

Department of Physics, Hokkaido University

(Received December 17, 1953)

The anomalous magnetic moments of nucleons are calculated, based upon the same standpoint as was employed in our previous work on pion-nucleon scatterings, where the Rarita-Schwinger field representing nucleon isobars is introduced and the charge independent interaction is assumed. In this paper we assume as the interaction Hamiltonian with external magnetic field the interaction energy of normal magnetic moments of nucleons and isobars with external magnetic field. Comparisons are made of our results with those of the covariant calculations in ps-ps theory, which clarify the validity of our non-relativistic calculations and it is shown that the usual covariant perturbation calculation would not be able to explain the anomalous magnetic moments of nucleons if the damping effect of nucleon pair creation and annihilation processes could be well taken into accounts. As a conclusion, the anomalous moments seem to be explained for the first time by the introduction of nucleon isobars, which seem to have to be described as if they were purely the elementary particles of spin 3/2.

§ 1. Introduction and summary

Recent experimental results on, especially, pion-nucleon scatterings and gamma-pion production seem to indicate the possible existence of isobar states of nucleons. Several phenomenological approaches have been done, assuming ps-ps and ps-pv couplings, for pion-nucleon scatterings[1],[2] and for the evaluation of isobar effect to the static nuclear potential[3]. In all these calculations, it is assumed that the nucleons in isobar states can be described by the Rarita-Schwinger theory[4] for spin 3/2 particles and an effective interaction Hamiltonian is introduced besides the usual ones, which permits the virtual transitions of nucleons to or from their isobar states. In our previous work on pion-nucleon scatterings[2] a satisfactory agreement with experimental data was obtained in the static approximations for nucleons by the lowest order perturbation calculations and the various adjustable parameters were determined quite reasonably and rather sharply.

The possibility that the anomalous magnetic moments of nucleons might as well be explained by the presence of nucleon isobars was indicated by one of the authors[5]. Recently T. Hamada[6] investigated this problem by performing the covariant calculations based upon the same footing as those mentioned above. However, no definite conclusions were obtained mainly because of extremely strong divergences due to the introdution of higher spin field. One way to overcome this difficulty is to treat nucleons at rest. Thus we calculated in this paper the anomalous magnetic moments of nucleons, starting from the same standpoint as was employed by our previous investigation on pion-nucleon scatterings[2], i. e. in the static approximations for nucleons by the lowest order perturbation calculations.

The aim of our present calculation is, therefore, to draw some definite conclusions on the possibility indicated by one of us[5] by calculating the anomalous magnetic moments of nucleons quite consistent with the explanation of pion-nucleon scattering data in terms of isobar assumption, i. e. by carring out the the calculations in the same approximations as in our previous work on pion-nucleon scatterings[2] and using the values of various adjustable parameters determined there.

In § 2, the calculations are made in the above mentioned line. We calculate the self-energy of a nucleon in a weak magnetic field in the approximations mentioned above, using the cut-off procedure. We use the same interaction Hamiltonian between nucleon and meson fields as those employed in our previous work[2], which consists of two parts, one of which is, in ps-ps theory, the sum of equivalent ps-pv term and the meson pair term and the other of which is the effective interaction term permitting the virtual transitions of nucleons to or from their excited states. Beside these terms we assume the interaction terms with constant external magnetic field; one part of them is the usual expression representing the interaction of meson current with magnetic field and the other terms represent the interactions of normal magnetic moments of nucleons and their isobars with the external magnetic field. The latter expression contains the knowledge concerning the intrinsic magnetic moment of an elementary particle of spin $3/2$. According to Belinfante[7],[*] the intrinsic magnteic moment of spin $3/2$ particle of charge e and mass m is given by

$$(e/3mc) M, \qquad\qquad (1)$$

where M is its spin angular momentum. Whereas one of us assumed the value just three times as large as the above one in his earlier investigation.[5] Therefore Belinfante[7],[8] concluded that the suggestion indicated by one of us[5] that the anomalous magnetic moments of nucleons might well be explained by the introduction of nucleon isobars would be misleading. As is shown later, the satisfactory results are obtained only if we assume the value gievn by (1). Therefore it is concluded that the anomalous magnetic moments of nucleons may well be eplained by the introduction of nucleon isobars, where the nucleon isobars must be described as if they were purely the elementary particles of spin $3/2$.

In § 3, the covariant calculations are performed in the same way as the previous work by Hamada[6]. The only new point of our calculation is the modification of the propagation function of the Rarita-Schwinger field[11],[9]; we replace its some nonrelativistic approximation quite consistently with calculations in the previous section, and it gives rise to suppress the divergences at the same time. But the results depend on the cut-off momentum too sensible to couclude something after all, though the divergences are only logarithmic.

Therefore, in the final section, discussions are made only about the results of the

calculations of § 2. If we cut off all the integrals at nucleon rest mass, the best values for the anomalous moments are 2.22 for proton and -2.20 for neutron, (in units of nucleon magneton), using the values of the parameters determined by pion-nucleon scatterings.[2] The effect of the characteristic pair term is shown to be very small, if we take the accounts of the damping effect of the pair term[10]. Finally by comparing the results of our calculations with those of the covariant calculations in ps-ps theory by Nakabayashi and Sato[11], it is shown that our static approximation for nucleons seems to give the correct meson contribution, whereas it gives only a small value for nucleon contribution; thus our model gives too large proton moments compared with neutron moments. It is also shown that the very large favorable fourth order correction obtained by Nakabayashi and Sato[11] in the covariant method seems to be due to the neglection of the damping of the pair term and it seems to us that the usual covariant perturbation calculation in ps-ps theory would not be able to explain the anomalous moments of nucleons. As a summary we can conclude that the anomalous magnetic moments might for the first time be well explained by the introduction of nucleon isobars, thus supporting the previous suggestion by one of us[5].

§ 2.　The non-covariant calculation

We calculate the lowest orders self-energy of a nucleon in a weak magnetic field to obtain the anomalous magnetic moments. The interaction Hamiltonian is divided into two parts;

$$H'(x) = H_A(x) + H_B(x), \qquad (2)$$

where $H_A(x)$ is the interacton Hamiltonian among the nucleon, its isobar and meson fields, assuming the ps-ps coupling, and $H_B(x)$ is the one with an external constant magnetic field. We take the same form for $H_A(x)$ as in our previous work on pion-nucleon scattering[2]; namely

$$
\begin{aligned}
&H_A(x) = H^{(1)}(x) + H^{(2)}(x) + H^{(3)}(x) + H^{(4)}(x), \\
&H^{(1)}(x) = (f^2/2m)\,\overline{\psi}\psi\phi_\alpha{}^2 \rightarrow (f^2/2m)\,\phi^2\delta(x), \\
&H^{(2)} = (f/2m)\,\overline{\psi}i\gamma_5\gamma_\nu\tau_\alpha\psi\,\frac{\partial\phi_\alpha}{\partial x_\nu} \rightarrow (g/\mu)\,\sigma\nabla\phi_\alpha\tau_\alpha\delta(x)\,;\ g=f(\mu/2m), \\
&H^{(3)}(x) = (G/\mu)\,\overline{\Psi}_\nu T_\alpha\psi\,\frac{\partial\phi_\alpha}{\partial x_\nu} \rightarrow (G/\mu)\,\Pi\nabla\phi_\alpha T_\alpha\delta(x), \\
&H^{(4)}(x) = H^{(3)}(x)^\dagger,
\end{aligned}
\qquad (3)
$$

where the units $\hbar = c = 1$ are used $H^{(1)}(x)$ is the meson pair term and $H^{(2)}(x)$ is the equivalent pv coupling term, which are the results of the Dyson transformation[12] of the linear ps coupling term. f and g are the ps and pv coupling constants and m and μ are the masses of a nucleon and a meson respectively. ψ, Ψ_ν and ϕ_α represent the nucleon, its isobar, and meson fields respectively. The symbol $\rightarrow$ means to the approximation of nucleons at rest. The expression of Π is obtained from the four

independent positive energy solutions of the Rarita-Schwinger equation at rest, and is given by

$$\Pi = \begin{vmatrix} e_2 & 0 \\ 0 & e_1 \\ \dfrac{1}{\sqrt{3}}e_1 & \dfrac{2}{3}e_3 \\ -\sqrt{\dfrac{2}{3}}e_3 & \dfrac{1}{\sqrt{3}}e_2 \end{vmatrix},$$

where

$$e_1 = \{1/\sqrt{2} \quad i/\sqrt{2} \quad 0\}, \quad e_2 = \{1/\sqrt{2} \quad -i/\sqrt{2} \quad 0\}, \quad e_3 = \{0 \ 0 \ 1\}.$$

Under the assumption that a nucleon is at rest $H_B(x)$ is the interaction energy with a constant magnetic field of the normal moments of a nucleon and its isobar and the meson current and is given by

$$H_B(x) = H^{(5)}(x) + H^{(6)}(x) + H^{(7)}(x),$$
$$H^{(5)}(x) = -H_z \sigma_z q(e/2m)\delta(x), \quad H^{(6)}(x) = -H_z \sum{}'_z Q\epsilon (e/2m)\delta(x), \qquad (4)$$
$$H^{(7)}(x) = eH_z\phi_1(x\partial/\partial y - y\partial/\partial x)\phi_2,$$

where q and Q are the charge operators of the nucleon and its isobar and σ_z and $\sum{}'_z$ are the z components of the spin operators of the nucleon and its isobar respectively, which have the expressions

$$q = \frac{1}{2}(1+\tau_3) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad Q = \begin{vmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{vmatrix},$$

$$\sigma_z = \begin{vmatrix} 1 & 0 \\ 0 & -1 \end{vmatrix}, \quad \sum{}'_z = \begin{vmatrix} 3 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -3 \end{vmatrix}.$$

ϵ determines the gyromagnetic ratio of nucleon in its isobar state and is given by $1/3$ according to (1) if the nucleon isobar is purely an elementary particle of spin $3/2$ as was mentioned in the introduction. On the other hand ϵ must be much larger than $1/3$, if the nucleon isobar corresponds to some excited state of the surrounding meson cloud around the nucleon as is predicted, for example, by strong coupling theory, because the additional angular momentum h of an isobar is due to the meson cloud and the magnetic moment will correspondingly be much larger. Therefore we introduce here phenomenologically a

parameter ϵ and will determine the value of ϵ so as to fit the empirical moments.

The lowest order self-energy of a nucleon, which gives the anomalous magnetic moment, is as is well known, given by

$$\Delta E = \sum_{I_1}' \sum_{I_2}' \frac{\langle 0 \mid H' \mid I_2 \rangle \langle I_2 \mid H' \mid I_1 \rangle \langle I_1 \mid H' \mid 0 \rangle}{(E-E_1)(E-E_2)} , \tag{5}$$

where $\mid 0 \rangle$ is the initial state. $\mid I_1 \rangle$ and $\mid I_2 \rangle$ are the intermediate states, and E_1, E_2 are the energies in these states. The symbol $\sum'$ means to take the principal values. The diagrams, which contribute to the lowest order calculations are shown in Fig. 1.
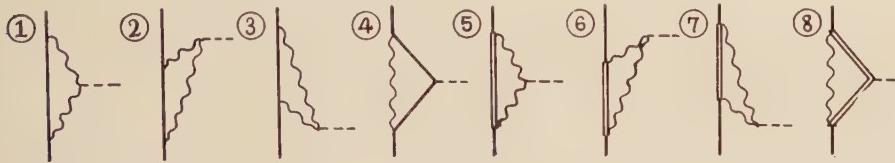
Fig. 1. These are all graphs, which contribute to the lowest order anomalous magnetic moment, where the various lines have the usual meanings.

The meson fields are expanded as usual;

$$\phi_\alpha(x) = \frac{1}{\sqrt{V}} \sum_k \frac{1}{\sqrt{2\omega_k}} (a_k^\alpha + a_{-k}^{\alpha *}) e^{ikx}, \ \omega_k = \sqrt{k^2 + \mu^2} . \tag{6}$$

The self-energies corresponding to these diagrams are easily calculated;

$$\Delta E(1) = \left(\frac{g}{\mu}\right)^2 \frac{eH_z}{4} \frac{1}{(2\pi)^6} \int dx \int d\mathbf{k} d\mathbf{k}' \frac{(\sigma\mathbf{k}')(\sigma\mathbf{k})}{\omega^2\omega'^2} \left(k_x' \frac{\partial}{\partial k_y} - k_y' \frac{\partial}{\partial k_x}\right)$$
$$\times (\tau_2\tau_1 - \tau_1\tau_2) e^{i(k-k')x}$$

$$= -\left(\frac{e}{2m}\right) H_z \sigma_z \left(\frac{g^2}{4\pi}\right) \frac{4}{3\pi} \left(\frac{m}{\mu}\right)^2 \tau_3 \frac{1}{m} \int_0^K \frac{k^4 dk}{\omega^4} ,$$

$$\Delta E(4) = -\left(\frac{g}{\mu}\right)^2 \frac{H_z}{2} \left(\frac{e}{2m}\right) \tau_\alpha q \tau_\alpha \frac{1}{(2\pi)^3} \int d\mathbf{k} \frac{(\sigma\mathbf{k})\sigma_z(\sigma\mathbf{k})}{\omega^3}$$

$$= \left(\frac{e}{2m}\right) H_z \sigma_z \left(\frac{g^2}{4\pi}\right) \frac{1}{6\pi} \left(\frac{m}{\mu}\right)^2 (3-\tau_3) \frac{1}{m^2} \int_0^K \frac{k^4 dk}{\omega^3} ,$$

$$\Delta E(5) = \left(\frac{G}{\mu}\right)^2 \frac{eH_z}{4} \frac{1}{(2\pi)^6} \int dx \int d\mathbf{k} d\mathbf{k}' \frac{(H^*\mathbf{k}')(H\mathbf{k})(T_2^*T_1 - T_1^*T_2)}{\omega\omega'(\Delta\epsilon+\omega)(\Delta\epsilon+\omega')}$$
$$\times \left(k'_x \frac{\partial}{\partial k_y} - k'_y \frac{\partial}{\partial k_x}\right) e^{i(k-k')x}$$

$$= -\left(\frac{e}{2m}\right) H_z \sigma_z \left(\frac{G^2}{4\pi}\right) \frac{8}{27\pi} \left(\frac{m}{\mu}\right)^2 \tau_3 \frac{1}{m} \int_0^K \frac{k^4 dk}{\omega^2(\Delta\epsilon+\omega)^2} , \tag{7}$$

$$\varDelta E(6) = -\left(\frac{G}{\mu}\right)^2 \frac{cH_z}{(2\pi)^6} \int d\!\!r \int dk dk' \; \frac{1}{\omega\omega'} \left\{ \frac{(\varPi^*k)(\varPi k) T_2^* T_1}{(\omega+\omega')(\varDelta\epsilon+\omega')} \right.$$

$$\left. + \frac{(\varPi^*k)(\varPi k) T_1^* T_2}{(\omega+\omega')(\varDelta\epsilon+\omega)} \right\} \left(k'_y \frac{\partial}{\partial k'_x} - k'_x \frac{\partial}{\partial k'_y} \right) e^{-i(k+k')x}$$

$$= -\left(\frac{e}{2m}\right) H_z \sigma_z \left(\frac{G^2}{4\pi}\right) \frac{4}{27\pi} \left(\frac{m}{\mu}\right)^2 \tau_3 \frac{1}{m} \int_0^K \frac{k^4 dk}{\omega^3(\varDelta\epsilon+\omega)} \quad,$$

$$\varDelta E(8) = -\left(\frac{G}{\mu}\right)^2 \frac{H_z}{2} \epsilon \left(\frac{e}{2m}\right) \frac{1}{(2\pi)^3} \int dk \frac{(\varPi^*k) \sum_z (\varPi k) T_\alpha^* Q T_\alpha}{\omega(\varDelta\epsilon+\omega)^2} .$$

$$= -\left(\frac{e}{2m}\right) H_z \sigma_z \left(\frac{G^2}{4\pi}\right) \frac{4\epsilon}{27\pi} \left(\frac{m}{\mu}\right)^2 (3+5\tau_3) \frac{1}{m^2} \int_0^K \frac{k^4 dk}{\omega(\varDelta\epsilon+\omega)^2} \quad,$$

$$\varDelta E(2) = \varDelta E(3) = \frac{1}{2} \varDelta E(1), \quad \varDelta E(7) = \varDelta E(6),$$

where K is the cut-off momentum and $\varDelta\epsilon$ is the excitation energy of isobar.   In the course of these calculations the following relations* are used;

$$T_2^* T_1 - T_1^* T_2 = \frac{4}{3} i\tau_3, \quad T_\alpha^* Q T_\alpha = \frac{2}{3}(3+5\tau_3),$$

$$(\varPi^*k')(\varPi k) = (k'k) - \frac{1}{3}(\sigma k')(\sigma k),$$

$$(\varPi^*k) \sum_z (\varPi k) \rightarrow 2/9 \cdot \sigma_z k^2.$$

In the last equation we use the fact that $k_i k_j$ is equal to $1/3 \cdot k^2 \delta_{ij}$ in the integrand, because the other factors of the integrand involve only k^2.   Then the anomalous moments are given as follows;

the usual meson contribution;

$$\delta \mathfrak{M}_{u.m.} = \frac{4}{3\pi} \left(\frac{g^2}{4\pi}\right) \left(\frac{m}{\mu}\right)^2 \tau_3 \frac{1}{m} \int_0^K \frac{k^4 dk}{\omega^4} ,$$

the usual nucleon contribution;

$$\delta \mathfrak{M}_{u.n.} = -\frac{1}{6\pi} \left(\frac{g^2}{4\pi}\right) \left(\frac{m}{\mu}\right)^2 (3-\tau_3) \frac{1}{m^2} \int_0^K \frac{k^4 dk}{\omega^3} ,$$

the isobar meson contribution;   (8)

$$\delta \mathfrak{M}_{i.m.} = \frac{8}{27\pi} \left(\frac{G^2}{4\pi}\right) \left(\frac{m}{\mu}\right)^2 \tau_3 \left[\frac{1}{m} \int_0^K \frac{k^4 dk}{\omega^2(\varDelta\epsilon+\omega)^2} + \frac{1}{m} \int_0^K \frac{k^4 dk}{\omega_3(\varDelta\epsilon+\omega)} \right],$$

---

* Details are given in another papers, for example, T. Matsumoto, T. Hamada and M. Sugawara, Prog. Theor. Phys. **10** (1953), 199.

the isobar nucleon contribution ;

$$\delta \mathfrak{M}_{i.n.} = \frac{4\epsilon}{27\pi} \left(\frac{G^2}{4\pi} \right) \left(\frac{m}{\mu} \right)^2 (3 + 5\tau_3) \frac{1}{m^2} \int_0^K \frac{k^4 dk}{\omega (\Delta \epsilon + \omega)^2} .$$

§ 3. The covariant calculation

About the covariant method of calculations the detailed investigation has been given by T. Hamada[i]. The interaction energy between the anomalous magnetic moment and the constant magnetic field is identified with the expression ;

$$-\delta \mathfrak{M} \sum_{\mu < \nu} F_{\mu\nu}(x) \overline{\psi}(x) \sigma_{\mu\nu} \psi(x), \quad \sigma_{\mu\nu} = (1/2i)[\gamma_\mu, \gamma_\nu] , \tag{9}$$

in the effective Hamiltonian

$$H_F(x) = -\frac{1}{2} \int dx_1 dx_2 T \{ H^e(x), \; H^i(x_1), \; H^i(x_2) \}, \tag{10}$$

where $H^e(x)$ is an interaction Hamiltonian with external electromagnetic field and $H^i(x)$ is the one between meson and nucleon. To obtain the additional effects due to the isobar, the interaction Hamiltonians are taken as

$$H^e(x) = e A_\mu^e(x) [-i \overline{\Psi}_\lambda(x) Q \gamma_\mu \Psi_\lambda(x) + (\varphi_1 \partial \varphi_2 / \partial x_\mu - \varphi_2 \partial \varphi_1 / \partial x_\mu)] ,$$

$$H^i(x) = (G/\mu) [\overline{\Psi}_\mu T_i \psi + \overline{\psi} T_i^* \Psi_\mu] \partial \varphi_i / \partial x_\mu . \tag{11}$$

The one nucleon portion of (10), diagrams of which are given in Fig. 2, gives

$$H_n(x) = \frac{-ieG^2}{12\mu^2} A_\mu^e(x) \int dx_1 dx_2(x_1) \overline{\psi} \mathfrak{S}_F^{\nu\lambda}(x - x_1) \gamma_\mu$$

$$\times \mathfrak{S}_F^{\lambda\sigma}(x_2 - x)(3 + 5\tau_3) \psi(x_2) \Delta_F^{\nu\sigma}(x_1 - x_2), \tag{12}$$

for the isobar nucleon contribution, and

$$H_m(x) = \frac{ieG^2}{6\mu^2} A_\mu^e(x) \int dx_1 dx_2 \overline{\psi}(x_1) \mathfrak{S}_F^{\lambda\nu}(x_1 - x_2) \tau_3 \psi(x_2)$$

$$\times \{ \Delta_F^\nu(x_1 - x) \Delta_F^{\lambda\mu}(x - x_2) + \Delta_F^\lambda(x - x_3) \Delta_F^{\nu\mu}(x_1 - x) \}, \tag{13}$$

for the isobar meson contribution, where $\mathfrak{S}_F^{\mu\nu}(x)$ and $\Delta_F(x)$ are the propagation functions of the Rarita-Schwinger fields and the meson field respectively, and are given by[1),9)]

$$\mathfrak{S}_F^{\mu\nu}(x) = \left[(\gamma \partial - M) \left\{ \delta_{\mu\nu} - \frac{1}{3} \gamma_\mu \gamma_\nu + \frac{1}{3M}(\gamma_\mu \gamma_\nu - \gamma_\mu \gamma_\nu) - \frac{2}{3M^2} \gamma_\mu \gamma_\nu \right\} \right.$$

$$\left. + \frac{2}{3M^2}(\Box - M^2) \{ (\gamma_\mu \partial_\nu - \gamma_\nu \partial_\mu) + (\gamma \partial - M) \gamma_\mu \gamma_\nu \} \right] \Delta_F(x)$$

$$= \left[\left\{ \delta_{\mu\nu} - \frac{1}{3} \gamma_\mu \gamma_\nu - \frac{1}{3M}(\gamma_\mu \partial_\nu - \gamma_\nu \partial_\mu) - \frac{2}{3M^2} \gamma_\mu \gamma_\nu \right\} (\gamma \partial - M) \right.$$

$$-\frac{2}{3M^2}\{-(\gamma_\mu\partial_\nu-\gamma_\nu\partial_\mu)+\gamma_\mu\gamma_\nu(\gamma\partial-M)\}(\square-M^2)\Big]\varDelta_F(x),$$

$$\varDelta_F(x)=\frac{2i}{(2\pi)^4}\int\frac{e^{ikx}}{k^2+\mu^2}d^4k,\qquad\varDelta_F^\mu(x)=\frac{\partial}{\partial x_\mu}\varDelta_F(x),$$

where $M=m+\varDelta\epsilon$ is the mass of the isobar.

As $\mathfrak{S}_F^{\mu\nu}(x)$ has strong singularities in the relativistic region, we modify the function to decrease these singularities. As $\mathfrak{S}_F^{\nu\mu}(x)$ has two equivalent expressions (see (14)), then we can obtain two different forms of the non-relativistic approximate expressions;
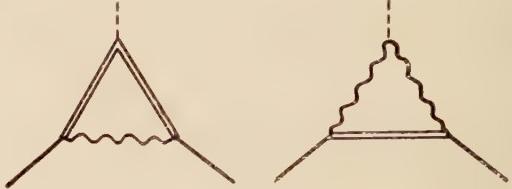
Fig. 2. The diagrams contributing to the isobar effects in the usual covariant calculation up to the lowest order perturbation.

$$\mathfrak{S}_F^{\mu\nu}(x)\approx(\gamma\partial-M)(\partial_{\mu\nu}-1/3\cdot\gamma_\mu\gamma_\nu)\varDelta_F(x),\tag{15}$$

$$\mathfrak{S}_F^{\mu\nu}(x)\approx(\partial_{\mu\nu}-1/3\cdot\gamma_\mu\gamma_\nu)(\gamma\partial-M)\varDelta_F(x).\tag{16}$$

These approximations mean a certain neglection of recoils, and seem to be consistent, because, as we can see later, they give the same results in spite of their different forms, indicating that the non-relativistic calculations in the previous section can have definite meaning.

3. 1   Isobar nucleon contribution

The Fourier transform of (12) is*

$$H_n^{(1)}(x)=\frac{-2eG^2}{3\mu^2(2\pi)^4}A_\mu^e(x)\int d^4k\frac{\bar\psi(x)[i\gamma(p'-k)-M](\partial_{\nu\lambda}-1/3\cdot\gamma_\nu\gamma_\lambda)\gamma_\mu}{[(p'-k)^2+M^2][(p-k)^2+M^2]}$$
$$\times\frac{[i\gamma(p-k)-M](\partial_{\lambda\sigma}-1/3\cdot\gamma_\lambda\gamma_\sigma)k_\nu k_\sigma(3+5\tau_3)\psi(x)}{(k^2+\mu^2)},\tag{17}$$

or

$$H_n^{(2)}(x)=\frac{-2eG^2}{3\mu^2(2\pi)^4}A_\mu^e(x)\int d^4k\frac{\bar\psi(x)(\partial_{\nu\lambda}-1/3\cdot\gamma_\nu\gamma_\lambda)[i\gamma(p'-k)-M]\gamma_\mu}{[(p'-k)^2+M^2][(p-k)^2+M^2]}$$
$$\times\frac{(\partial_{\lambda\sigma}-1/3\cdot\gamma_\lambda\gamma_\sigma)[i\gamma(p-k)-M]k_\nu k_\sigma(3+5\tau_3)\psi(x)}{(k^2+\mu^2)},\tag{18}$$

according to (15) or (16) respectively. The procedure of the calculation is straightforward. At the first the denominator of (17) or (18) is symmetrized by using the Feynman's formula;

*   We understand here $\psi(x)$ and $\bar\psi(x)$ as $\psi(p)e^{ipx}$ and $\bar\psi(p')e^{-ip'x}$ respectively.

$$\frac{1}{abc} = 2\int_0^1 x\,dx \int_0^1 dy \frac{1}{[a(1-x) + bxy + cx(1-y)]^3}.$$

$$\left. \begin{aligned} A_\mu^e(x)\bar{\psi}(x)p_\mu\psi(x) &\approx -1/2 \sum_{\mu<\nu} F_{\mu\nu}(x)\bar{\psi}(x)\sigma_{\mu\nu}\psi(n), \\ A_\mu^e(x)\bar{\psi}(x)\gamma_\mu\bar{\tau}_\nu \Delta p_\nu\psi(x) &\approx -\sum_{\mu<\nu} F_{\mu\nu}(x)\bar{\psi}(x)\sigma_{\mu\nu}\psi(x) \end{aligned} \right\} \tag{19}$$

are extracted, which give the interaction of the magnetic moments with the constant magnetic field.

The results are given by*

$$\left. \begin{aligned} H_n^{(\Gamma)}(x) &\approx -\frac{4eG^2 A_\mu^e(x)}{3\mu^2(2\pi)^4} \int_0^1 x\,dx \int_0^1 dy\, \{A/[k^2+\Lambda]^3\}\,\bar{\psi}(3+5\tau_3)\psi, \\ H_n^{(1)}(x) &\approx -\frac{4eG^2 A_\mu^e(x)}{3\mu^2(2\pi)^4} \int_0^1 x\,dx \int_0^1 dy\, \{B/[k^2+\Lambda]^3\}\,\bar{\psi}(3+5\tau_3)\psi, \end{aligned} \right\} \tag{20}$$

where

$$A = -\frac{iM}{9}(3k^2 - 2m^2x^2)xy\gamma_\mu\bar{\tau}_\nu\Delta p_\nu + \frac{i}{9}(7m + 5M)x(3k^2 - 2m^2x^2)p_\mu$$

$$- \frac{4i}{9}(m+M)(k^2 - 2m^2x^2)p_\mu - \frac{10}{9}im\,x^2(2k^2 - m^2x^2)p_\mu - \frac{4i}{9}$$

$$(m+M)^2 m x^2 p_\mu,$$

$$B = A + \frac{2i}{9}M xy(3k^2 - 2m^2x^2)\gamma_\mu\bar{\tau}_\nu\Delta p_\nu - \frac{2i}{9}M(3k^2 - 2m^2x^2)xp_\mu,$$

$$\Lambda = \mu^2 + (M - m^2 - \mu^2)x + m^2x^2.$$

In the course of the above calculations the following relations are used;

$$\gamma p\psi(x) = im\psi(x), \quad \bar{\psi}(x)\gamma p' = im\bar{\psi}(x),$$

$$p\Delta p A_\mu^e(x) = (\Delta p)^2 A_\mu^e(x) = \Delta p_\mu A_\mu^e(x) = 0, \quad p^2 = p'^2 = -m^2.$$

As B is effectively equal to A on account of (19), the two modifications of the Green's function of the Rarita-Schwinger field are equivalent in the calculation of magnetic moments.

By comparing (20) with (9), the isobar nucleon contribution to the magnetic moments given by

$$\partial \mathfrak{M}_m = \left(\frac{G^2}{4\pi}\right)\left(\frac{m}{\mu}\right)^2 \frac{1}{27\pi}\left(\frac{e}{2m}\right)\int_0^1 dx\left\{\left[-4\left(1+\frac{M}{m}\right)x + 3\left(7 + 4\frac{M}{m}\right)x^2 - 20x^3\right]\right.$$

$$\times\left[2\log(K/m) - \log(\Lambda/m^2) - \frac{1}{2}\right] + \left[2\left(1+\frac{M}{m}\right)\left(1-\frac{M}{m}\right)\frac{x^3}{\Lambda/m^2}\right.$$

---

* By $\approx$ we mean the equivalency in the calculation of the magnetic moments.

$$-\left(7+4\frac{M}{m}\right)\frac{x^4}{\Lambda/m^2}+5\frac{x^5}{\Lambda/m^2}\right]\right\}(3+5\tau_3),$$

where K is a cut-off momentum and we have performed the integrations to ∞ for the finite integrals and have used cut-off procedure only for divergent integrals.

3. 2 Meson contribution

The Fourier transform of (13) is

$$H_m^{(1)}(x)$$
$$=\frac{4ieG^2\cdot V_\mu(x)}{3\mu^2(2\pi)^4}\int d^4k\,\bar{\psi}(x)\,(i\gamma k-M)(\partial_{\lambda\nu}-1/3\cdot\gamma_\lambda\gamma_\nu)(p'-k)_\nu(p+p'-2k)_\mu\tau_3\psi(x),$$
$$\frac{}{(k^2+M^2)[(p'-k)^2+\mu^2][(p-k^2)+\mu^2]},$$

or   (22)

$$H_m^{(2)}(x)$$
$$=\frac{4ieG^2\cdot V_\mu(x)}{3\mu^2(2\pi)^4}\int d^4k\,\bar{\psi}(x)\,(\partial_{\lambda\nu}-1/3\cdot\gamma_\lambda\gamma_\nu)(i\gamma k-M)(p'-k)_\nu(p+p'-2k)_\mu\tau_3\psi(x),$$
$$\frac{}{(k^2+M^2)[(p'-k)^2+\mu^2][(p-k)^2+\mu^2]},$$

corresponding to the modification (15) and (16) respectively. The calculations are quite the same way as before and in this case also, $H_m^{(1)}(x)$ and $H_m^{(2)}(x)$ give the same contributions to the magnetic moments.

The result is given by

$$H_m(x)\approx\frac{8ieG^2 A_\mu^e(x)}{3\mu^2(2\pi)^4}\int_0^1(1-x)dx\,\{C/[k^2+\Lambda]^3\}\,\bar{\psi}\tau_3\psi,\tag{23}$$

where

$$C=-\frac{k^2}{6}(m+M)\gamma_\mu\gamma_\nu\Delta p_\nu+\frac{2}{3}\left[-k^2(mx+M)(1-x)+k^2m(1-x)^2\right.$$
$$-k^2(mx+M)(1-x)+2m^2(mx+M)(1-x)^3\right]p_\mu.$$

Then the isobar meson contribution to the anomalous moments is

$$\delta\mathfrak{M}_m=-\left(\frac{G^2}{4\pi}\right)\left(\frac{m}{\mu}\right)^2\frac{1}{3\pi}\left(\frac{e}{2m}\right)\tau_3\int dx\left[\left\{-\frac{1}{6}\left(1+\frac{M}{m}\right)-\frac{5}{6}\left(1+\frac{M}{m}\right)x\right.\right.$$
$$+\left(\frac{7}{3}+\frac{M}{m}\right)x^2-\frac{4}{3}x^2\right\}\left\{2\log(K/m)-\log(\Lambda/m^2)-\frac{1}{2}\right\}$$ (24)
$$+\left\{\frac{1}{3}\left(1+\frac{M}{m}\right)\frac{x^3}{\Lambda/m^2}-\frac{1}{3}\left(2+\frac{M}{m}\right)\frac{x^4}{\Lambda/m^2}+\frac{1}{3}\frac{x^5}{\Lambda/m^2}\right\}\right],$$

where we used again the cut-off procedure only for the divergent integrals.

As can be seen from the two expressions (21) and (24), the values of $\delta\mathfrak{M}_n$ and $\delta\mathfrak{M}_m$ depend very much upon the cut-off momentum K and we cannot decide even the signs of these two expressions uniquely. This is due to the fact that we did not

use the cut-off procedure for finite integrals and thus our above calculations are not consistent with the cut-off procedure which means the cut-off of all the integrations with the virtual meson momentum. However, if we want to do, we must deal with much more complicated mathematical expressions, which indicates that the cut off procedure, an essentially nonrelativistic procedure, is by no means suitable one in the covariant calculations. As we are based upon rather questionable background, it seems of no value to go into mathematical details further and we will not discuss the results of this section at all in the following section.

§ 4. The numerical results and the discussions

In this section we discuss only about the results of the nonrelativistic calculation.

First of a all, we give here the values of the various parameters, which are determined so as to attain the best agreements with experimental results in our previous work[2] :

$$g^2/4\pi \equiv \alpha_g = 0.06 \sim 0.07, \quad G^2/4\pi \equiv \alpha_G = 0.1 \pm 0.05 ,$$

$$\Delta\epsilon = 2.3 \sim 2.4 \, \mu, \tag{25}$$

the damping coefficient for the pair term $\approx 1/6$.

Table I. The numerical results of the nonrelativistic calculations of the anomalous moments with $a_g = (g^2/4\pi)$ and $a_G = (G^2/4\pi)$ in units of nucleon magneton. $\delta\mathfrak{M}_n$ and $\delta\mathfrak{M}_m$ mean the nucleon and meson contributions, respectively.

	$\delta\mathfrak{M}_P$		$\delta\mathfrak{M}_N$	
	usual	isobar	usual	isobar
$\delta\mathfrak{M}_n$	$-\ 2.067\ a_g$	$+10.534\ a_G$	$-\ 4.134\ a_g$	$-2.633\ a_G$
$\delta\mathfrak{M}_m$	$+25.969\ a_g$	$+\ 4.704\ a_G$	$-25.969\ a_g$	$-4.704\ a_G$
total	$+23.902\ a_g$	$+15.238\ a_G$	$-30.103\ a_g$	$-9.337\ a_G$
$a_g = 0.07$ $a_G = 0.1$	$+3.30$		-2.84	
$a_g = 0.06$ $a_G = 0.05$	$+2.22$		-2.20	

We can see readily from the above table that we can not assign to ϵ much larger values than $1/3$, because the larger ϵ makes the isobar nucleon contributions too large to fit the empirical moments. Thus it seems to us that the nucleon isobars must be described as if they were purely the elementary particles of spin $3/2$, which means that we must introduce into the present theory the new degrees of freedom corresponding to the nucleon isobar.

Next it is very interesting to compare our above results with those of the covariant calculations of Nakabayasi and Sato[11] in *ps-ps* theory, which are given in Table II. By

comparing our usual contributions with their second order contributions and considering that α_f. $177\alpha_n$, we can conclude that our nonrelativistic calculations give the correct meson contributions, whereas the nucleon contributions have much smaller magnitude owing to the neglection of nucleon current contributions. As the usual nucleon contributions are negative both for proton and neutron, the suppression of these contributions means the increase of proton moment relative to neutron moment. The fact that our calculated values of proton moment are always larger in magnitude than neutron one is due to this circumstance and it is expected that the more correct method of calculation will explain this discrepancy of our results with empirical values.

Table II. Contributions of normal interaction to anomalous moments up
to the fourth order in symmetrical ps-ps theory with $a_f = f^2/4\pi$.
$\delta\mathfrak{M}_n$ and $\delta\mathfrak{M}_m$ mean the nucleon and meson contributions, respectively.

	$\delta\mathfrak{M}_P$		$\delta\mathfrak{M}_N$	
	second order	fourth order	second order	fourth order
$\delta\mathfrak{M}_n$	$-0.075\ a_f$	$+0.0306\ a_f^2$	$-0.151\ a_f$	$-0.015\ a_f^2$
$\delta\mathfrak{M}_m$	$-0.111\ a_f$	$+0.0025\ a_f^2$	$-0.111\ a_f$	$-0.0025\ a_f^2$
total	$+0.036\ a_f$	$+0.033\ a_f^2$	$-0.262\ a_f$	$-0.018\ a_f^2$

As we can suppose that the fourth order contributions of covariant calculations are mainly due to the characteristic pair term, because the damping effect for pair term is not there taken into account, we investigate here the contributions of the pair term in (3) to the anomalous magnetic moments. In our static approximations, it is easily shown that the meson contributions vanish exactly and the nucleon contributions have good sign for proton contrary with the second order nucleon contributions, and have the expression

$$\delta\mathfrak{M}_{\text{pair}} = (f^2/4\pi)^2_{\text{pair}} \frac{3q}{16\pi^2} \left[\frac{1}{m} \int_0^\kappa \frac{k^2 dk}{\omega^2} \right]^2, \tag{26}$$

where the approximation $\omega + \omega' \to 2\sqrt{\omega\omega'}$ was made in the energy denominator, which gives the over-estimated value for the integral. The value of the above expression is $0.01q(f^2/4\pi)^2_{\text{pair}}$, which is the same order of magnitude as the fourth order nucleon contribution for proton in Table II. The small magnitude of the other fourth order effects in Table II. corresponds to the exact vanishment of the pair term meson contributions in our case. On the other hand, if we choose $(f^2/4\pi)_{\text{pair}} \approx 1.5$ as is required by the pion-nucleon scattering data[a], the value of $\delta\mathfrak{M}_{\text{pair}}$ is $0.026q$ and is very small compared with the usual second order contributions. Thus we are led to the conclusion that the usual covariant perturbation calculation in ps-ps theory would not be able to explain the anomalous magnetic moments of nucleons, if the damping of the processes of nucleon-pair creation and annihilation could be well taken into accounts.

Our calculated values in Table I. are rather too large in magnitude. But they depend upon the cut-off momentum K and they can be made much smaller if we choose K smaller than nucleon mass, which is more reasonable, because the cut-off momentum must be smaller than the nucleon mass. As a summary we can conclude that the anomalous magnetic moments of nucleons might be at the first time well explained by the introduction of nucleon isobars into the theory by some way, thus justifying the previous suggestion by one of us.[5]

The authors would like to express their gratitude to Messrs. Homada, Matsumoto and Shyono for their many valuable discussions.

References

1) S. Minami, T. Nakano, K. Nishijima, H. Okonogi and E. Yamada, Prog. Theor. Phys. 8(1952), 531.
2) A. Kanazawa and M. Sugawara, Prog. Theor. Phys. 10 (1953), 399.
3) T. Matsumoto, T. Hamada and M. Sugawara, Prog. Theor. Phys. 10 (1953), 199.
   T. Matsumoto and M. Sugawara, Prog. Theor. Phys. (to be published).
4) Rarita and J. Schwinger, Phys. Rev. 56 (1941), 61.
5) M. Sugawara, Prog. Theor. Phys. 8 (1952), 549.
6) T. Hamada, Prog. Theor. Phys. 10 (1953), 309.
7) F. J. Belinfante, Phys. Rev. 92 (1953), 997.
8) F. J. Belinfante, Phys. Rev. 92 (1953), 994.
9) Y. Takahashi and H. Umezawa, Prog. Theor. Phys. 9 (1953), 14.
10) K. A. Brueckner, M. Gell-Mann and M. Goldberger, Phys. Rev. 90 (1953), 476.
11) K. Nakabayasi and I. Sato, Sci. Rep. Tohoku Univ. Ser. 1. 34 (1950), 169.
12) F. J. Dyson, Phys. Rev. 86 (1952), 802.

On the Bose-Einstein Liquid Model for Liquid Helium, III*

— Further Considerations regarding Vapour Pressures of He^3-He^4 Mixtures —

Ziro MIKURA

*The Research Institute for Iron, Steel and Other Metals,
Tohoku University, Sendai*

Discontinuities at the λ-point in the temperature derivatives of vapour pressures of a He^3-He^4 mixture, which were predicted by deBoer and Gorter, have been thought not to be found in the existing experimental results. Re-examination of these data, however, reveals that the partial vapour pressure of He^3 has a break when it is plotted as a function of temperature. Both the deBoer - Gorter - Taconis and the "modified Bose-Einstein liquid" theories agree with experiment with respect to this point, in contrast to the Heer and Daunt theory. Below the λ-temperature, the "modified B. E. liquid" theory gives the partial pressure of He^3 in qualitatively good agreement with experiment, but above this temperature neither theory agrees with experiment even qualitatively. The temperatures at which experimental these discontinuities occur, are however always considerably higher than the λ-temperatures expected from the experiment on superflow in mixtures.

§ 1. Introduction

In the previous two papers[1],[2] of this series (referred to hereinafter as Papers (I) and (II) respectively), the "modified Bose-Einstein liquid" theory was developed and applied to various problems on He^3-He^4 mixtures and those on pure liquid He^4 under high pressures. The vapour pressure in the He^3-He^4 mixture, which has already been partly discussed in Paper (I), will be further discussed in this paper, since it is a problem in purely static equilibrium, independent of the somewhat ambiguous dynamical behaviours of He^3 and He^4 atoms in the mixture, and consequently the study of it enables one to select among various possible ways, a most justifiable one to construct the thermodynamic functions of He^3-He^4 liquid mixtures. deBoer and Gorter[3] used the thermodynamic functions of a mixture of ideal classical liquids under Taconis' assumption, and Heer and Daunt[4] used those of a mixture of a nondegenerate ideal Fermi-Dirac gas (liquid He^3) and a degenerate ideal Bose-Einstein gas (liquid He^4) in respective potential wells. The present author presented a theory, in which Heer and Daunt's ideal B. E. gas was replaced by a "modified B. E. gas" with an excitation energy proportional to the number density of He^4 particles, whose individual mass is chosen as 8.8 times as large as the mass of a He^4 atom, independently of He^3 concentration. In § 2, the relation between the vapour

---

pressure and temperature is considered. Among the above mentioned theories, those of deBoer and Gorter and the author predict breaks in the vapour pressure versus temperature curve at the λ-temperature, while Heer and Daunt's theory does not. In the experimental data of Weinstock *et al*[5]. and Sommers[6], in which no breaks have been thought to be found, breaks are shown to exist when the partial vapour pressure of He[3] is plotted as a function of temperature*. It is found, however, that the breaks in these experimental plots are always less marked than the theoretical ones, and are located at considerably higher temperatures than those expected from the experimental results of superflow in mixtures. The general features of p_3-T curves below the λ-temperatures in the "modified B. E. liquid" theory agree with those of experimental plots, though the agreement above these temperatures is not good. The results of the deBoer-Gorter-Taconis theory does not seem to agree with experiment even qualitatively. In § 3, the vapour pressure versus He[3] concentration relations at constant temperature are discussed.

§ 2. Partial vapour pressure of He[3] versus temperature relation

From purely thermodynamic considerations assuming the transition of the second order, deBoer and Gorter[7] showed that the temperature derivatives of the total and partial vapour pressures in the He[3]-He[4] mixture should have discontinuities at the λ-temperature. It has been thought that in the experiments of both Weinstock *et al.* and Sommers no such discontinuities were observed. The author evaluated, from their data of total vapour pressure, the values of the partial pressure of He[3], p_3, and plotted them as a function of temperature. Breaks in p_3-T plots are then clearly found in the four following cases : Sommers' 9.49 and 13.0% solutions and Weinstock *et al.*'s 20.3 and 25·5% solutions. In the case of Sommers' more dilute solutions, no breaks are found. The reason why the breaks appear clearly in the p_3-T plots, but not in the p-T plots seems obvious : in solutions with relatively small He[3] concentrations, the relative amount of discontinuity in dp_3/dT is large and that in dp_4/dT is small, but, as the absolute value of dp_4/dT is much larger than that of dp_3/dT, the relative amount of discontinuity in $dp/dT (=dp_3/dT + dp_4/dT)$ becomes small.

The partial vapour pressure of He[3], p_3, can be derived from experimental values of total vapour pressure, p, by applying

$$p_3 = p . X_v$$

and the empirical equation

$$X_v = \Delta p/p . [1 + 0.080 \{T - 1.100 + 0.600 (\Delta p/p)^3\}^2] ,$$

where X_v is the mole concentration of He[3] in vapour, and Δp the total vapour pressure of a solution minus the vapour pressure of pure He[4] at the same temperature. This formula was presented by Sommers[6] in order to represent his experimental results for

---

* In a lecture given in Kyoto on Sept. 22, 1953, Gorter reported that in very recent Leiden observations a discontinuity is found in the temperature derivative of the vapour to liquid concentration ratio at the λ-point.

$0 \leq X_r \leq 1.00$ and in the temperature range $1.3° \leq T \leq 2.2°K$. In this paper this formula is used up to temperatures a little higher than $2.2°K$.
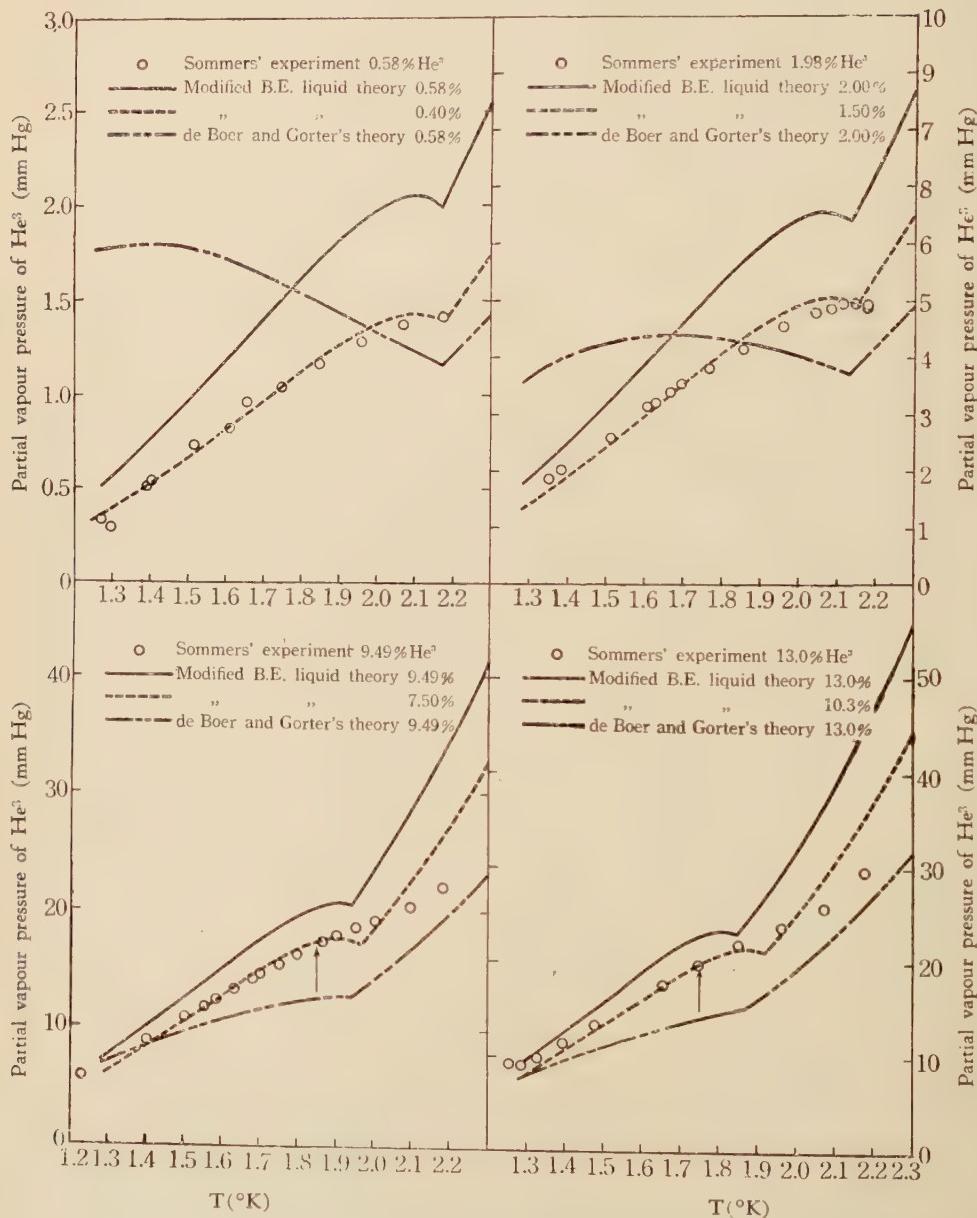
Fig. 1. The partial vapour pressure of He³ plotted as a function of temperature. Experimental points are derived from the unsmoothed data of Sommers[6] for the total vapour pressure. The arrows indicate the expected λ-temperatures. (*Note added in proof*: The positions of the arrows are erroneous. The expected λ-temperatures should nearly coincide with the break points of the uppermost (solid) curves.)

In Fig. 1, the values of p_3 derived from Sommers' unsmoothed data are plotted as a function of T. The plots for 9.49 and 13.0% solutions are clearly seen to have breaks, though the number of experimental points near the break-temperatures are too small to locate them definitely. Thus, the prediction by deBoer and Gorter[7] may be said to be experimentally confirmed. The break for a 9.49% solution is situated at a temperature higher than 2.0° K, and that for 13.0% solution at a temperature higher than 1.95°K, as is seen in the figure. A remarkable fact is that no indications of breaks are found at the λ-temperatures expected from the experiment of Abraham *et al.*[5] on superflow through the superleak immersed in solutions, and that the break-temperatures found in the present plots are considerably higher than the expected λ-temperatures, 1.94° and 1.86°K (indicated by arrows in the figure)*. Fig. 2 shows the similar plots for 20.3 and 25.5% solutions of Weinstock *et al.* As they did not present numerical values of p, the points in the

(a) ○ Weinstock et al's experiment 20.3% He³
   ——— Modified B. E. liquid theory 20.3%
   - - - -   ,,   ,,   17.0%
   – – – deBoer and Gorter's theory 20.3%

(b) ○ Weinstock et al's experiment 25.5% He³
   ——— Modified B. E. liquid theory 25.5%
   - - - - deBoer and Gorter's theory 25.5%
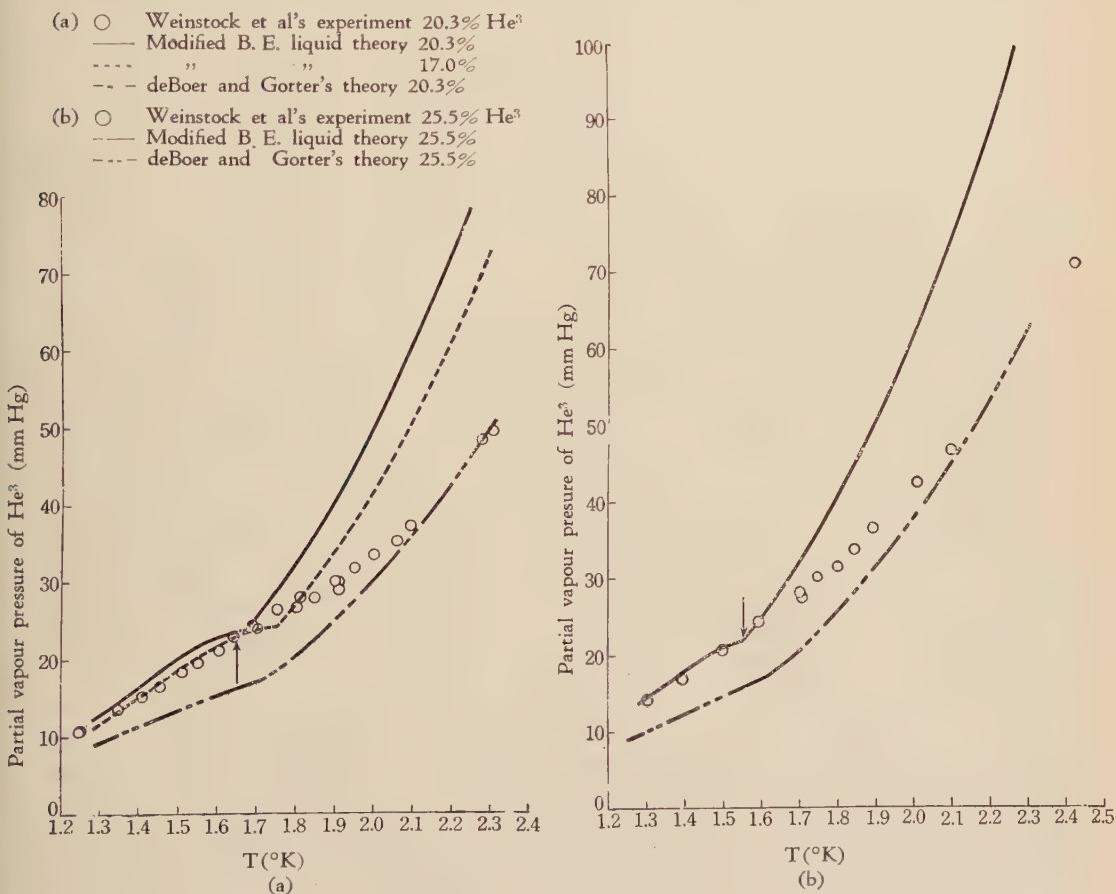

Fig. 2. The partial vapour pressure of He³ plotted as a function of temperature. Experimental points are derived from the log p versus $1/T$ plots of Weinstock *et al*[5]. (a) 20.3% He³, (b) 25.5% He³. The arrows indicate the expected λ-temperatures.

---

* Note added in proof: The positions of the arrows in Fig. 1 are erroneous.

figure are derived by reading the experimental points in their $\log p$ versus $1/T$ graph. Breaks are also found in these cases, at about $1.85°$ K for the 20.3 % solution and at about $1.75°$ K for the 25.5 % solution. They are again considerably higher than the expected λ-temperatures, 1.65' and 1.55' K respectively (indicated by arrows in the figure).

Now the results are compared with theories. Since experimental results clearly show jumps in dp_s/dT at λ-temperatures, those theories alone are considered, which are in accord with experiment with respect to this point. In Figs. 1 and 2, the results from the deBoer-Gorter-Taconis theory are shown by dot-dashed curves, and those from the "modified B. E. liquid' theory by solid curves. The formulae for vapour pressures in both theories, which are used to draw these curves, are found in Paper (I) (eqs. $(34)-(43)$). The dotted curves indicate the results of calculation from the "modified B. E. liquid" theory when the He^3 concentrations in solutions are taken somewhat smaller than those reported by the original authors. The agreement below the theoretical λ-temperatures can be made satisfactory by this adjustment of the values of concentration. Deviations above the λ-temperatures, however, still remain. It should further be noted that the breaks in the empirical p_s-T plots are still located at temperatures considerably higher than the theoretical λ-points for the solutions of adjusted concentrations.

The above discrepancies between the transition temperatures obtainable from the two kinds of measurement, i.e. the vapour pressure and the rate of superflow, appear to be too large to be ascribed to experimental errors in the determination of He^3 concentration in either of these experiments. If one assumes the determination of concentration in Abraham et al'.s superflow experiment not seriously erroneous, the 20.3 % solution of Weinstock et $al.$ should have contained only about 13 % He^3, and the 25.5 % solution about 17 % He^3. These values are still lower than those of concentrations which must be assumed, in the "modified B. E. liquid" theory, in order to obtain the agreement with experiment below the λ-temperatures. (See Fig. 2) If the discrepancies were not caused by the experimental errors in the determination of concentration of the original gas, one is led to assume that the vapour on the solution does not exhibit the true equilibrium pressure corresponding to the liquid concentration, say, on account of the decrease of He^3 concentration on the liquid or in the film on the wall. This possibility might be tested by an experiment, in which the vapour pressure and the rate of superflow through the super-leak immersed in a solution are simultaneously observed.

The quite different features of theoretical and experimental plots above the λ-temperature appear to suggest that the "modified B. E. liquid" theory is no longer valid in this temperature region. This suggestion seems to be supported also by the following fact: the specific heat of pure He^4 evaluated from this theory is about seven times smaller than the experimentally obtained value just above the λ-temperature, though it is in satisfactory agreement with experiment below this temperature. The low values of the theoretical specific heat above the λ-temperature is obviously caused by the assumption of small particle number, i. e. $1/\nu$ times the number of atoms ($\nu=8.8$, c. f. Paper (I)). It is noticeable that, if one takes the particle number equal to the number of atoms, the theoretical specific heat above the λ-temperature is of the same order of magnitude as that

experimentally obtained, though assuming different values for ν above and below the λ-temperature introduces complications in the explanation of the λ-transition as the B. E. gas condensation.

§ 3. Vapour pressure versus He³ concentration relation

There have been so far no experimental studies on vapour pressures of solutions containing more than 25.5 % He³. From the theoretical point of view, the investigations of these cases seem to be interesting, in particular those of the case of pure liquid He³ containing small amount of He⁴. The two theories discussed in the preceding section predict quite different features in this concentration region. This is clearly shown in Fig. 3, in which the total vapour pressure is plotted as functions of He³ concentration. The results of calculation from the Heer and Daunt "ideal B. E. liquid" theory are also
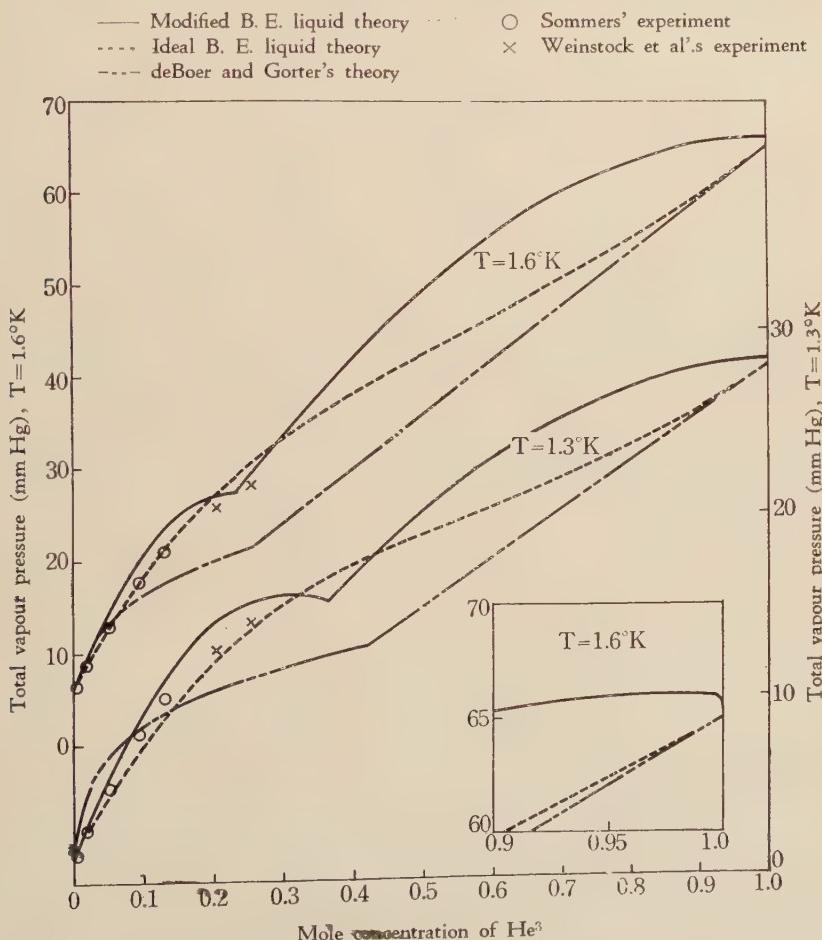

Fig. 3. The total vapour pressures of He³-He⁴ mixtures plotted as a function of He³ concentration. The insert shows in large scale the feature in the neighbourhood of pure He³.

included, together with the experimental points in the low concentration region. The difference between the deBoer-Gorter-Taconis and the " modified B. E. liquid " theories are drastic in the neighbourhood of 100 % He³. (See the insert in Fig. 3). The peculiar feature of the curve, which is seen, in the case of the " modified B. E. liquid " theory, in the neighbourhood of pure He³, originates clearly from the decrease of the number of He⁴ particles in the liquid state due to clustering   (c. f. eq. (39) in Paper (I)). The validity of this assumption of small particle number in liquid state is open to question above the λ-temperature as was discussed in the preceding section, and consequently so is also this characteristic deviation from Raoult's law at extremely high He³ concentrations. The solution of the problem is left to future experiments.

§ 4.   Conclusion

In concluding this series of papers, it may be said that the " modified B. E. liquid " model of liquid He¹ with the mass factor, ν, and the excitation energy, $\varDelta$, satisfying the relations*

$$\varDelta = \varDelta_0 \frac{n/\nu}{n_0/\nu_0} \left(\frac{M_4}{M_4^0} \right)^{1/3},$$ (1)

$$\nu = M_4/m_4 = \nu_0 + 0.15 \, (p - p_0),$$ (2)

represents reasonably well the real properties of liquid He⁴, pure and diluted with He³, under high pressures as well as under its saturated vapour pressures, at least below the λ-temperatures.** It should be remarked, however, that eq. (1) has been set forth in order to reproduce as well as possible the empirical relation between λ-temperature and He³ concentration, and that whether it is correct or not depends strongly on the accuracy of this experiment. Consequently possible future revisions of eqs. (1) and (2) are reserved until more precise experiments on superflow in solution will appear, since, on account of the scattering of the experimental points in existing observations, one cannot consider these results as most conclusive.

References

1)   Z. Mikura, Prog. Theor. Phys. **11** (1954), 25.
2)   Z. Mikura, Prog. Theor. Phys. **11** (1954), 207.
3)   J. deBoer and C. J. Gorter, Physica **16** (1950), 225.
4)   C. V. Heer and J. G. Daunt, Phys. Rev. **81** (1951), 448.
5)   B. Weinstock, D. W. Osborne, and B. M. Abraham, Phys. Rev. **77** (1950), 400.
6)   H. S. Sommers, Phys. Rev. **88** (1952), 113.
7)   J. deBoer and C. J. Gorter, Physica **18** (152), 565.
8)   B. M. Abraham, B. Weinstock, and D. W. Osborne, Phys. Rev. **76** (1949), 864.

* Here n is the number density of He⁴ atoms, M_4 the particle mass, m_4 the mass of a He⁴ atom, and p the pressure in atmosphere. The suffix zero means the pure He⁴. cf. Paper (II).
** The properties bearing on the equation of states, such as the thermal expansion and compressibility, are excluded, because they depend on the properties of the potential well, which is assumed rigid in the present model.

On Gauge Invariance and the Structure of Elementary Particles

Yasushi TAKAHASHI*

Physical Institute, Nagoya University

Non-gauge terms inherent in current field theory are discussed. An identity which holds, independently of the definition of integration, between the non-gauge term and the vacuum expectation value of the energy-momentum tensor of vacuum particles is derived and its physical meaning is discussed.

§ 1. Introduction

As is well-known, the current theory of elementary particles has many unresolved difficulties. According to the renormalization theory of Tomonaga and Schwinger, some of these difficulties can be avoided, and so far as quantum electrodynamics is concerned, we can obtain an answer which might have physical significance. In spite of this fact, current quantum electrodynamics has two characteristic inconsistencies, namely, the two invariance requirements, which are satisfied in the initial formulation of the theory, gauge invariance and relativistic covariance, are not fulfilled at the end of certain types of calculation. These results have been discussed by many authors in terms of problems of the self-energy of the photon[1] and the self-stress of the electron.[2]

These circumstances lead us to suspect even those results which were obtained by renormalization and which were believed to have physical significance.

The second problem, however, has been reduced to that of the diverging self-energy of the electron by Pais-Epstein and others,[2] and it is now fairly clear that the problem is one of appearance only and that if one is able to obtain a finite electron self-energy, then the difficulty will be resolved.

On the other hand, in spite of many efforts, the nature of the first problem, the self-energy of the photon, is still not so clear and the problem has remained a hopeless one.

Heretofore, most of the discussions of the self-energy of the photon have been to calculate it in a straightfoward way and regard it as the problem of the definition of integration.[1] Thus it would be of interest to attempt to reduce the problem to a simpler one from a different point of view.

Since the problem of the vacuum polarization always involves subtraction of the vacuum which involves finding a suitable definition of the vacuum at the initial stage of the calculation, we shall settle the problem by considering the relation between the vacuum

---

* Now at the University of Rochester, Rochester, N. Y., U.S.A.

polarization and the density of vacuum particles. In fact, Umezawa-Kawabe[3] developed an interesting discussion along this line from a correspondence-theoretical standpoint.

In this paper, we would like to show that the self-energy of the photon is related in a simple way to the energy-momentum tensor of the vacuum particle and discuss the meaning of this relation. Furthermore, in the same way, we will derive and discuss an identity between the non-gauge matrix element of the two gamma decay of π^0 and the energy-momentum tensor of the vacuum particles.

It is believed that the method of analysis discussed may provide us with a powerful tool for the evaluation of the non-equivalent term of this process, but this part of the problem shall be considered in the future.

§2. Connection between the self-energy of the photon and the vacuum expectation value of the energy-momentum tensor of the particle

The current induced by the interaction between the electromagnetic field and the vacuum particle* is, in the 2nd order perturbation calculation,

$$\delta j_\mu(x) = \frac{(-i)^2}{2!} \int dx' \langle P(j_\mu(x), H(x')) \rangle_{\text{one photon}}$$

$$\equiv i\frac{e^2}{4} \int dx' K_{\mu\nu}(x-x') A_\nu(x') dx', \qquad (2\cdot1)$$

where

$$H(x) = -j_\lambda(x) A_\lambda(x), \qquad (2\cdot2)$$

$$K_{\mu\nu}(x-x') = S_p \{ S_F(x-x') \gamma_\mu S_F(x'-x) \gamma_\nu \}. \qquad (2\cdot3)$$

In order to separate the self-energy of the photon from $(2\cdot1)$, it is convenient to introduce Fourier integral transforms as follows :

$$\left. \begin{array}{l} K_{\mu\nu}(x-x') \equiv \int K_{\mu\nu}(k) e^{ik(x-x')} dk \\ A_\nu(x') \equiv \int A_\nu(p) e^{ipx'} dp. \end{array} \right\} \qquad (2\cdot4)$$

On substituting $(2\cdot4)$ into $(2\cdot1)$, we obtain

$$\delta j_\mu(x) = \frac{ie^2}{4} \int dx' \int dk \int dp K_{\mu\nu}(k) A_\nu(p) e^{ikx} e^{i(p-k)x'}$$

$$= \frac{ie^2}{4} (2\pi)^4 \int dk K_{\mu\nu}(k) A_\nu(k) e^{ikx}. \qquad (2\cdot5)$$

Now, only $K_{\mu\nu}(k)$ independent of k** enables us to replace $(2\cdot5)$ with

---

* We use the units $\hbar = c = 1$ throughout this paper. We assume for the time being that the vacuum particle has spin $\frac{1}{2}$.

** In this case we can write $K_{\mu\nu} \equiv K \delta_{\mu\nu}$ with a constant K.

$$\partial j_\mu(x) = \frac{ic^2}{4}(2\pi)^2 K_{\mu\nu} A_\nu(x) \tag{2.6}$$

$$= \frac{ic^2}{4}(2\pi)^2 K A_\mu(x) \tag{2.7}$$

and to obtain the self-energy of the photon. In this case, we have from (2.4)

$$K_{\mu\nu}(x-x') = (2\pi)^4 \delta_{\mu\nu} K \delta(x-x')$$

and

$$(2\pi)^4 K = \frac{1}{4} \int K_{\lambda\lambda}(x-x') dx'. \tag{2.8}$$

Thus we can separate the self-energy of the photon*.

Let us now transform (2.8) and derive an useful identity.

By use of (A.1), we have

$$\int K_{\lambda\lambda}(x-x') dx' = \int dx' S_p \{ S_F(x-x')\gamma_\lambda S_F(x'-x)\gamma_\lambda \}$$

$$= 2i(x-x')_\lambda S_p \{ \gamma_\lambda S_F(x-x') \}_{x \to x'}$$

$$= 8i(x-x')_\lambda \frac{\partial}{\partial x_\lambda} \Delta_F(x-x')|_{x \to x'}$$

$$= 8i \left[\frac{1}{\kappa} \frac{\partial}{\partial \kappa} \{ \kappa^2 \Delta_F(x-x') \} - 4\Delta_F(x-x') \right]_{x \to x'}. \tag{2.9}$$

On the other hand, the vacuum expectation value for the trace of the energy-momentum tensor is

$$\langle T_{\lambda\lambda} \rangle_0 = 2\kappa^2 \Delta_F(x-x')|_{x \to x'}, \tag{2.10}**$$

and so we obtain by substituting (2.10) into (2.9)

$$\int K_{\lambda\lambda}(x-x') dx' = 4i \left[\frac{1}{\kappa} \frac{\partial}{\partial \kappa} \langle T_{\lambda\lambda} \rangle_0 - \frac{4}{\kappa^2} \langle T_{\lambda\lambda} \rangle_0 \right] \tag{2.11}$$

and

$$\partial j_\mu(x) = -\frac{c^2}{4} \left\{ \frac{1}{\kappa} \frac{\partial}{\partial \kappa} \langle T_{\lambda\lambda} \rangle_0 - \frac{4}{\kappa^2} \langle T_{\lambda\lambda} \rangle_0 \right\} A_\mu(x). \tag{2.12}$$

In this way, we can reduce the problem of the self-energy of the photon to the $\langle T_{\lambda\lambda} \rangle_0$, *independently of the definition of integral.*

---

* The separation is possible without Fourier expansion, i.e., if we expand $A_\nu(x')$ into Taylor series $A_\nu(x') = A_\nu(x) + (x'-x)_\lambda \frac{\partial}{\partial x_\lambda} A_\nu(x) + \cdots$ and substitute this into (2.1). Then, the second term vanishes because of the Lorentz condition and the third term is the charge renormalization term.

** J. Schwinger, Phys. Rev. **75** (1949), 651.
The vacuum expectation value for the energy-momentum tensor of the total system is the same as the one for the vacuum particle only.

In the beginning of this section we assumed that the vacuum particle has spin $\frac{1}{2}$, but this restriction is in fact unnecessary, i.e., we can easily verify the identity $(2 \cdot 12)$ for spin 0 and 1, and in these cases

$$\langle T_{\lambda\lambda} \rangle_0 = -n \kappa^2 J_F(0),$$ $(2 \cdot 13)$

with

$$n = \begin{cases} 1 & \text{for spin } 0, \\ -2 & \text{for spin } \frac{1}{2}, \\ 3 & \text{for spin } 1. \end{cases}$$

We note from the identities of the spin 0, $\frac{1}{2}$ and 1 particles the following properties :
(i)   The self-energy of the photon as well as the vacuum energies of the respective particles are linear in n, i.e., $\partial j_\mu(x)$ is proportional to n.   This fact may be considered as the field-theoretical generalization of the discussion given by Umezawa and Kawabe[3] based upon the correspondence-theoretical point of view.   The success of mixed field theory in the problem of self-energy of the photon is also based on this relation.
(ii)   Upon carrying out the integration of $J_F(0)$ and substituting the results into $(2 \cdot 13)$, we have

$$\langle T_{\lambda\lambda} \rangle_0 = -n \frac{\kappa^2}{4\pi^2} \lim_{K \to \infty} \left(K^2 - \kappa^2 \left(\log \frac{2K}{\kappa} - \frac{1}{2} \right) \right),$$ $(2 \cdot 14)$

and therefore

$$\partial j_\mu(x) = -\frac{ne^2}{8\pi^2} \lim_{K \to \infty} \left(K^2 - \frac{1}{2}\kappa^2 \right) A_\mu(x),$$ $(2 \cdot 15)$

which is in agreement with the value calculated in a straight-forward manner by several authors.[1]   Furthermore we observe the following remarkable fact : since $\langle T_{\lambda\lambda} \rangle_0$ is a physical- quantity, in spite of its mathematical behavior, it must have the form, from dimensional considerations,

$$\langle T_{\lambda\lambda} \rangle_0 = a\, n\, \kappa^4$$

with a dimensionless constant a, thus if we substitute $(2 \cdot 16)$ into $(1 \cdot 12)$, we have

$$\partial j_\mu(x) = 0,$$

i.e., the self-energy of the photon vanishes. *We can get the gauge invariant result only when we regard $\langle T_{\lambda\lambda} \rangle_0$ not as a mathematical quantity but as a physical one.*
The cut-off momentum K in $(2 \cdot 14)$ is not a physical quantity in the truest sense of the word, but a temporary cut-off factor, which must be inevitable because of the divergence inherent in current theory, and therefore cannot contribute to the dimension. The physical meaning of this point will be discussed in detail in section 4.

§ 3. On the non-gauge matrix element of the process $\pi^0 \to 2\gamma$

We examine now the relation between $\langle T_{\lambda\lambda}\rangle_0$ and the non-gauge matrix element of the γ-decay of neutral mesons. We can write

$$\langle \bar{\psi}\gamma_L\psi\rangle = \frac{(-i)^2}{2!}\int dx^1 \int dx^2 \langle P(\bar{\psi}(x)\gamma_L\psi(x), H(x^1), H(x^2))\rangle_{\text{two photon}}$$

$$= -\frac{e^2}{8}\int dx^1 \int dx^2 K_{\mu\nu L}(x, x^1, x^2) A_\mu(x^1) A_\nu(x^2), \qquad (3\cdot1)$$

where

$$H(x) = -j_\lambda(x)A_\lambda(x)$$

$$\gamma_L = \begin{cases} 1 & \text{for scalar meson with scalar coupling,} \\ \gamma_0\gamma_5 & \text{for pseudoscalar meson with pseudovector coupling,} \end{cases} \qquad (3\cdot2)$$

and

$$K_{\mu\nu L}(x, x^1, x^2) = S_p\{S_F(x-x^1)\gamma_\mu S_F(x^1-x^2)\gamma_\nu S_F(x^2-x)\gamma_L\}. \qquad (3\cdot3)$$

In order to separate the non-gauge term in $(3\cdot1)$, we adopt the procedure discussed in section 2.

(i) In the case of $\gamma_L = 1$, by noting the transformation property of the matrix element, we conclude that the non-gauge term has a form $A_\mu(x)A_\nu(x)$. Therefore, according to the method in section 2 we have

$$\langle \bar{\psi}(x)\psi(x)\rangle = -\frac{e^2}{32}\int K_{\lambda\lambda}(x, x^1, x^2)\,dx^1 dx^2 A_\mu(x)A_\mu(x). \qquad (3\cdot4)$$

(ii) In the case of $\gamma_L = \gamma_0\gamma_5$, considering the transformation property again we note that we cannot have $A_\mu(x)A_\nu(x)$ terms but can have $\varepsilon_{\mu\nu\lambda\sigma}A_\mu(x)F_{\nu\lambda}(x)$ terms only, where $\varepsilon_{\mu\nu\lambda\sigma}$ equals $+1$ and -1 for even and odd permutation of $(\mu\nu\lambda\sigma)$, respectively. Therefore, by introducing the Taylor expansion

$$\left.\begin{aligned} A_\mu(x^1) &= A_\mu(x) + (x^1-x)_\lambda \frac{\partial}{\partial x_\lambda}A_\mu(x) + \cdots, \\ A_L(x^2) &= A_\nu(x) + (x^2-x)_\lambda \frac{\partial}{\partial x_\lambda}A_\nu(x) + \cdots, \end{aligned}\right\} \qquad (3\cdot5)$$

$(3\cdot1)$ can be rewritten as

$$\langle \bar{\psi}(x)\gamma_0\gamma_5\psi(x)\rangle$$

$$= -\frac{e^2}{8}\int K_{\mu\nu\sigma}(x, x^1, x^2)\,dx^1 dx^2 A_\mu(x)A_\nu(x)$$

$$- \frac{e^2}{8}\int K_{\mu\nu\sigma}(x, x^1, x^2)(x^1-x)_\lambda dx^1 dx^2 A_\nu(x)\frac{\partial A_\mu(x)}{\partial x_\lambda}$$

$$-\frac{e^2}{8}\int K_{\mu\nu\sigma}(x, x^1, x^2)(x^2-x)_\lambda dx^1 dx^2 A_\mu(x)\frac{\partial A_\nu(x)}{\partial x_\lambda}$$

$$+\cdots. \tag{3.6}$$

The first term vanishes by virtue of relativistic covariance consideration and since we can write

$$\left.\begin{array}{l} \int K_{\mu\nu\sigma}(x, x^1, x^2)(x^1-x)_\lambda dx^1 dx^2 \equiv K^{(1)}\varepsilon_{\mu\nu\lambda\sigma}, \\[2mm] \int K_{\mu\nu\sigma}(x, x^1, x^2)(x^2-x)_\lambda dx^1 dx^2 \equiv K^{(2)}\varepsilon_{\mu\nu\lambda\sigma}, \end{array}\right\} \tag{3.7}$$

with a constant $K^{(1)}$ and $K^{(2)}$, we have

$$\langle \bar\psi(x)\gamma_\sigma\gamma_5\psi(x) \rangle_{\text{non-gauge}} = -\frac{e^2}{8}K^{(1)}\varepsilon_{\mu\nu\lambda\sigma}A_\nu(x)\frac{\partial A_\mu(x)}{\partial x_\lambda}$$

$$-\frac{e^2}{8}K^{(2)}\varepsilon_{\mu\nu\lambda\sigma}A_\mu(x)\frac{\partial A_\nu(x)}{\partial x_\lambda}$$

$$=\frac{e^2}{8}(K^{(1)}-K^{(2)})\varepsilon_{\mu\nu\lambda\sigma}A_\mu(x)\frac{\partial A_\nu(x)}{\partial x_\lambda}$$

$$=-\frac{e^2}{16}(K^{(1)}-K^{(2)})\varepsilon_{\mu\nu\lambda\sigma}A_\mu(x)F_{\nu\lambda}(x), \tag{3.8}$$

where

$$F_{\nu\lambda}(x)=\frac{\partial A_\lambda(x)}{\partial x_\nu}-\frac{\partial A_\nu(x)}{\partial x_\lambda}.$$

Thus we have to know

$$K^{(1)}=\frac{e^2}{24}\int \varepsilon_{\mu\nu\lambda\sigma}K_{\mu\nu\sigma}(x, x^1, x^2)(x^1-x)_\lambda dx^1 dx^2$$

and

$$\left.\begin{array}{l}\\[2mm]\end{array}\right\} \tag{3.9}$$

$$K^{(2)}=\frac{e^2}{24}\int \varepsilon_{\mu\nu\lambda\sigma}K_{\mu\nu\sigma}(x, x^1, x^2)(x^2-x^1)_\lambda dx^1 dx^2,$$

where the relations

and

$$\left.\begin{array}{l} \varepsilon_{\mu\nu\lambda\sigma}=-\varepsilon_{\nu\mu\lambda\sigma} \\[2mm] \varepsilon_{\mu\nu\lambda\sigma}\varepsilon_{\mu\nu\lambda\sigma}=24 \end{array}\right\} \tag{3.10}$$

have been used.

(i) First, we discuss the case of the scalar meson. By use of $(A\cdot 2)$, we have

$$4K\equiv \int K_{\lambda\lambda}(x, x^1, x^2) dx^1 dx^2$$

$$=\int S_p\{S_F(x-x^1)\gamma_\lambda S_F(x^1-x^2)\gamma_\lambda S_F(x^2-x)\} dx^1 dx^2 \tag{3.11}$$

$$=-2(x-x')_\lambda(x-x')_\lambda S_p S_F(x-x')|_{x\to x'} \tag{3.12}$$

$$=8\kappa(x-x')_\lambda(x-x')_\lambda \Delta_F(x-x')|_{x\to x'}. \tag{3.13}$$

Furthermore $(A \cdot 3)$, $(A \cdot 4)$ and $(2 \cdot 10)$ enable $(3 \cdot 13)$ to be rewritten as

$$4K = \frac{4}{\kappa^2}\left[\frac{\partial^2}{\partial u^2}(\kappa^2 \varDelta_F(0)) - \frac{1}{\kappa}\frac{\partial}{\partial \kappa}(\kappa^2 \varDelta_F(0)) - 4\kappa \frac{\partial}{\partial \kappa}\varDelta_F(0) \right]$$

$$= 2\left[\frac{1}{\kappa^2}\frac{\partial^2}{\partial \kappa^2} - \frac{5}{\kappa^3}\frac{\partial}{\partial \kappa} + \frac{8}{\kappa^4} \right]\langle T_{\lambda\lambda} \rangle_0. \tag{3.14}$$

Therefore from $(3 \cdot 4)$ we obtain

$$\langle \bar{\psi}(x)\psi'(x) \rangle_{\text{non-gauge}} = -\frac{e^2}{16}\left[\frac{1}{\kappa^2}\frac{\partial^2}{\partial \kappa^2} - \frac{5}{\kappa^3}\frac{\partial}{\partial \kappa} + \frac{8}{\kappa^4} \right]\langle T_{\lambda\lambda} \rangle_0. l_\mu(x) . l_\mu(x). \tag{3.15}$$

(ii) Second, we would like to discuss the case of pseudoscalar with pseudovector coupling. From $(3 \cdot 8)$

$$\langle \bar{\psi}(x)\gamma_\sigma\gamma_5\psi(x) \rangle_{\text{non-gauge}} = -\frac{e^2}{16}(K^{(1)} - K^{(2)})\varepsilon_{\mu\nu\lambda\sigma}A_\mu(x)F_{\nu\lambda}(x). \tag{3.16}$$

Now we have to evaluate $K^{(1)}$ and $K^{(2)}$:

$$K^{(1)} = \frac{1}{24}\varepsilon_{\mu\nu\lambda\sigma}\int K_{\mu\nu\lambda\sigma}(x, x^1, x^2)(x^1 - x)_\lambda dx^1 dx^2$$

$$= \frac{1}{24}\varepsilon_{\mu\nu\lambda\sigma}\int S_p\{S_F(x - x^1)\gamma_\mu S_F(x^1 - x^2)\gamma_\nu S_F(x^2 - x)\gamma_\sigma\gamma_5\}(x^1 - x)_\lambda dx^1 dx^2$$

$$= \frac{1}{24}\varepsilon_{\mu\nu\lambda\sigma}S_p(\gamma_\rho\gamma_\mu\gamma_{\rho 1}\gamma_\nu\gamma_{\rho 2}\gamma_\sigma\gamma_5)$$

$$\times \int \frac{\partial}{\partial x_\rho}\varDelta_F(x - x^1)\frac{\partial}{\partial x_{\rho 1}^1}\varDelta_F(x^1 - x^2)\frac{\partial}{\partial x_{\rho 2}^2}\varDelta_F(x^2 - x)\cdot(x^1 - x)_\lambda dx^1 dx^2$$

$$= \frac{\kappa^2}{6}\varepsilon_{\mu\nu.\sigma}\left[\varepsilon_{\mu\lambda\rho 2\sigma}\int \varDelta_F(x - x^1)\varDelta_F(x^1 - x^2)\frac{\partial}{\partial x_{\rho 2}^2}\varDelta_F(x^2 - x)\cdot(x^1 - x)_\lambda dx^1 dx^2 \right.$$

$$+ \varepsilon_{\mu\rho 1\nu\sigma}\int \varDelta_F(x - x^1)\frac{\partial}{\partial x_{\rho 1}^1}\varDelta_F(x^1 - x^2)\cdot\varDelta_F(x^2 - x)\cdot(x^1 - x)_\lambda dx^1 dx^2$$

$$\left. + \varepsilon_{\rho\mu\nu\sigma}\int \frac{\partial}{\partial x_\rho}\varDelta_F(x - x^1)\cdot\varDelta_F(x^1 - x^2)\varDelta_F(x^2 - x)\cdot(x^1 - x)_\lambda dx^1 dx^2 \right]$$

(unfortunately we must use the partial integral here)

$$= \frac{1}{48}\varepsilon_{\mu\nu\lambda\sigma}S_p(\gamma_\rho\gamma_\mu\gamma_{\rho 1}\gamma_\nu\gamma_{\rho 2}\gamma_\sigma\gamma_5 + \gamma_\rho\gamma_\mu\gamma_{\rho 2}\gamma_\nu\gamma_{\rho 1}\gamma_\sigma\gamma_5)$$

$$\times \int \frac{\partial}{\partial x_\rho}\varDelta_F(x - x^1)\frac{\partial}{\partial x_{\rho 1}^1}\varDelta_F(x^1 - x^2)\frac{\partial}{\partial x_{\rho 2}^2}\varDelta_F(x^2 - x)\cdot(x^1 - x)_\lambda dx^1 dx^2$$

$$+ \frac{\kappa^2}{6}\varepsilon_{\mu\nu\lambda\sigma}\varepsilon_{\rho\mu.\sigma}\int \frac{\partial}{\partial x_\rho}\varDelta_F(x - x^1)\cdot\varDelta_F(x^1 - x^2)\cdot\varDelta_F(x^2 - x)\cdot(x^1 - x)_\lambda dx^1 dx^2$$

$$= \frac{1}{6} (\varepsilon_{\rho\mu\rho_2\sigma} \hat{\partial}_{\nu\rho_1} - \varepsilon_{\rho\mu\nu\sigma} \hat{\partial}_{\rho_1\rho_2} + \varepsilon_{\rho\mu\rho_1\sigma} \hat{\partial}_{\nu\rho_2})$$

$$\times \int \frac{\partial}{\partial x_\rho} \varDelta_F(x-x^1) \cdot \frac{\partial}{\partial x'_{\rho_1}} \varDelta_F(x^1-x^2) \cdot \frac{\partial}{\partial x^2_{\rho_2}} \varDelta_F(x^2-x) \cdot (x^1-x)_\lambda \, dx' dx^2$$

$$+ \frac{1}{\kappa^2} \int \frac{\partial}{\partial x_\rho} \varDelta_F(x-x^1) \cdot \varDelta_F(x^1-x^2) \cdot \varDelta_F(x^2-x) \cdot (x^1-x)_\rho \, dx^1 dx^2$$

$$= -\frac{1}{3} \int \frac{\partial}{\partial x_\rho} \varDelta_F(x-x^1) \frac{\partial}{\partial x^1_\rho} \varDelta_F(x^1-x^2) \frac{\partial}{\partial x^2_\rho} \varDelta_F(x^2-x) \cdot (x^1-x)_\sigma \, dx^1 dx^2$$

$$- \frac{1}{3} \int \frac{\partial}{\partial x_\sigma} \varDelta_F(x-x^1) \frac{\partial}{\partial x^1_\rho} \varDelta_F(x^1-x^2) \cdot \frac{\partial}{\partial x^2_\rho} \varDelta_F(x^2-x) \cdot (x^1-x)_\sigma \, dx^1 dx^2$$

$$- \frac{1}{3} \int \frac{\partial}{\partial x_\rho} \varDelta_F(x-x^1) \frac{\partial}{\partial x^1_\sigma} \varDelta_F(x^1-x^2) \frac{\partial}{\partial x^2_\rho} \varDelta_F(x^2-x) \cdot (x^1-x)_\sigma dx^1 dx^2$$

$$+ \kappa^2 \int \frac{\partial}{\partial x_\sigma} \varDelta_F(x-x^1) \cdot \varDelta_F(x^1-x^2) \cdot \varDelta_F(x^2-x) \cdot (x^1-x)_\sigma \, dx^1 dx^2$$

$$= \int (x-x^1)_\sigma \frac{\partial}{\partial x_\sigma} \varDelta_F(x-x^1) \cdot \frac{\partial}{\partial x^1_\rho} \varDelta_F(x^1-x^2) \cdot \frac{\partial}{\partial x^2_\rho} \varDelta_F(x^2-x) \, dx^1 dx^2$$

$$+ 2 \int \frac{\partial}{\partial x_\rho} \varDelta_F(x-x^1) \frac{\partial}{\partial x^1_\rho} \varDelta_F(x^1-x^2) \cdot \varDelta_F(x^2-x) dx^1 dx^2$$

$$- \kappa^2 \int (x-x^1)_\sigma \frac{\partial}{\partial x_\sigma} \varDelta_F(x-x^1) \cdot \varDelta_F(x^1-x^2) \cdot \varDelta_F(x^2-x) \cdot dx^1 dx^2. \qquad (3 \cdot 17)$$

Similarly we have

$$K^{(2)} = - \int \frac{\partial}{\partial x_\rho} \varDelta_F(x-x^1) \cdot \frac{\partial}{\partial x^1_\rho} \varDelta_F(x^1-x^2) (x^2-x)_\sigma \frac{\partial}{\partial x^2_\sigma} \varDelta_F(x^2-x) dx^1 dx^2$$

$$- 2 \int \varDelta_F(x-x^1) \frac{\partial}{\partial x^1_\rho} \varDelta_F(x^1-x^2) \frac{\partial}{\partial x^2_\rho} \varDelta_F(x^2-x) dx^1 dx^2$$

$$+ \kappa^2 \int \varDelta_F(x-x^1) \cdot \varDelta_F(x^1-x^2) \cdot (x^2-x)_\sigma \frac{\partial}{\partial x^2_\sigma} \varDelta_F(x^2-x) \cdot dx^1 dx^2. \qquad (3 \cdot 18)$$

Due to $(A \cdot 5)$ and $(A \cdot 6)$, these can be written as

$$K^{(1)} = \frac{1}{\kappa^2} \iint \left[\frac{\partial}{\partial \kappa} (\kappa^2 \varDelta_F(x-x^1)) - 4\kappa \varDelta_F(x-x^1) \right] \left[i \frac{\partial}{\partial \kappa} (\kappa^2 \varDelta_F(x^1-x)) \right] dx^1$$

$$+ \frac{2i}{\kappa} \int \frac{\partial}{\partial \kappa} (\kappa^2 \varDelta_F(x-x^1)) \cdot \varDelta_F(x^1-x) dx^1$$

$$- \kappa \int \left[\frac{\partial}{\partial \kappa} (\kappa^2 \varDelta_F(x-x^1)) - 4\kappa \varDelta_F(x-x^1) \right] \left[\frac{i}{\kappa} \frac{\partial}{\partial \kappa} \varDelta_F(x'-x) \right] dx^1$$

$$= -\frac{1}{\kappa^2}\left[\frac{\partial^2}{\partial\kappa^2}(\kappa^2 \Delta_F(0)) - 9\kappa\frac{\partial}{\partial\kappa}\Delta_F(0) - 24\Delta_F(0)\right]$$

$$-\frac{1}{\kappa^2}\left[\frac{\partial^2}{\partial\kappa^2}(\kappa^2 \Delta_F(0)) - \kappa\frac{\partial}{\partial\kappa}\Delta_F(0) - 24\Delta_F(0)\right]$$

$$= -\frac{2}{\kappa^2}\left[\frac{\partial^2}{\partial\kappa^2}(\kappa^2 \Delta_F(0)) - 5\kappa\frac{\partial}{\partial\kappa}\Delta_F(0) - 24\Delta_F(0)\right], \tag{3.19}$$

and

$$K^{(2)} = -K^{(1)}.$$

Therefore we obtain

$$\langle \bar{\psi}(x)\gamma_\sigma\gamma_5\psi(x)\rangle_{\text{non-gauge}} = \frac{c^2}{4}\left[\frac{\partial^2}{\partial\kappa^2}(\kappa^2\Delta_F(0)) - 5\kappa\frac{\partial}{\partial\kappa}\Delta_F(0) - 24\Delta_F(0)\right]$$

$$\times \varepsilon_{\mu\nu\lambda\sigma}A_\mu(x)F_{\nu\lambda}(x)$$

$$= \frac{c^2}{8}\left[\frac{\partial^2}{\partial\kappa^2}\langle T_{\lambda\lambda}\rangle_0 - \frac{5}{\kappa}\frac{\partial}{\partial\kappa}\langle T_{\lambda\lambda}\rangle_0 + \frac{8}{\kappa^2}\langle T_{\lambda\lambda}\rangle_0\right]\varepsilon_{\mu\nu\lambda\sigma}A_\mu(x)F_{\nu\lambda}(x). \tag{3.20}$$

We can thus observe the fact that non-gauge terms of $\langle \bar{\psi}(x)\psi(x)\rangle$ and $\langle \bar{\psi}(x)\gamma_\sigma\gamma_5\psi(x)\rangle$ are proportional to

$$\left(\frac{\partial^2}{\partial\kappa^2} - \frac{5}{\kappa}\frac{\partial}{\partial\kappa} + \frac{8}{\kappa^2}\right)\langle T_{\lambda\lambda}\rangle_0. \tag{3.21}$$

If we substitute (2·14) into (3·15) and (3·20), then we have the same result as obtained by Fukuda and Miyamoto.[1] According to the dimensional consideration similar to section 2, (3·15) and (3·20) vanish and the gauge invariant results are obtained.

§ 4. Summary and discussion

It has been shown in section 2 and section 3 that the terms which violate the gauge invariance requirement, i.e., the self-energy of the photon and the matrix element of γ-decay process of π^0, have a simple connection with the vacuum expectation value of the energy-momentum tensor of the vacuum particles. However, these identities are valid only in the second order approximation of perturbation calculation and we do not know at present whether these can be extrapolated to higher order or not.*

The formula of Pais-Epstein used in the discussion of self-stress[1] or the identities (2·12), (3·15) and (3·20) in this paper suggest the possibility of showing the relation between any two physical quantities by using the operator $\partial/\partial\kappa$.** At any rate, the

---

* It seems at a glance very easy although tedious for us to prove these for higher order, but it may be not only tedious but also not so meaningful.

** If we had a formulation which included consistently the operator $\partial/\partial\kappa$, then (2·12), (3·15) and (3·20) might be obtained only with the gauge invariance requirement of the theory and the Pais-Epstein formula with the relativistic covariance.

relations $(2 \cdot 12)$, $(3 \cdot 15)$ and $(3 \cdot 20)$ and the Pais-Epstein formula implies that although the quantities to be subtracted by renormalization and vacuum subtraction, i.e., $\partial \kappa$ and $\langle T_{\lambda\lambda} \rangle_0$, diverge mathematically, they are quantities with a physical dimension, thus they must not be subtracted as a mere mathematical quantity. Physically speaking, this fact can be seen as follows: Letting $K = 1/a$, it can be seen immediately that a corresponds to the radius of the photon. Thus the procedure of differentiating with respect to the mass before carrying out the limiting process clearly contradicts the fundamental assumption of a point particle incorporated in current field theory. By carrying out the limiting process $a \to 0$ first, the physical dimension of $\langle T_{\lambda\lambda} \rangle_0$ depends only on κ and one is thereby able to obtain properly the vanishing self-energy of the photon and gauge invariant matrix element of $\langle \bar{\psi}(x)\psi(x) \rangle$ and $\langle \psi(x) \gamma_0 \gamma_5 \psi(x) \rangle$.

It is of special importance that the above discussion does not suggest the "definitions of integration." We have merely taken into account the fact that any physical quantities must have a proper dimension—not withstanding their mathematical behavior and one is able to employ this analysis if some convenient relation involving this quantity exists.

As can be shown easily, since processes including closed loops always contain $\partial/\partial \kappa$, the ambiguity of integration is due to the above reason.

Thus it can be seen that one is able to obtain relativistic covariant and gauge-invariant results, only if one considers these quantities which are to be subtracted out as physical quantities and not merely as mathematical ones.

Acknowledgement

Appendix

(1) *Derivation of formulae*

$$(\gamma \partial + \kappa) S_F(x - x') = 2i\delta(x - x')$$

and

$$(\gamma \partial + \kappa)(x - x')_\lambda S_F(x - x') = \gamma_\lambda S_F(x - x'),$$

therefore we obtain

$$(x - x')_\lambda S_F(x - x') = \frac{1}{2i} \int S_F(x - x'') \gamma_\lambda S_F(x'' - x') dx''. \qquad (A \cdot 1)$$

In the similar manner we have

$$(x - x')_\lambda (x - x')_\nu S_F(x - x')$$

$$= \frac{1}{(2i)^2} \int S_F(x-x'') \gamma_\lambda S_F(x''-x''') \gamma_\nu S_F(x'''-x') dx'' dx'''$$

$$+ \frac{1}{(2i)^2} \int S_F(x-x'') \gamma_\nu S_F(x''-x''') \gamma_\lambda S_F(x'''-x') dx'' dx'''. \tag{A·2}$$

Next the identities involving $\varDelta_F(x-x')$ are discussed.

$$(\Box-\kappa^2) \varDelta_F(x-x') = 2i\delta(x-x'),$$

$$\therefore \qquad (\Box-\kappa^2) \frac{\partial}{\partial\kappa} \varDelta_F(x-x') = 2\kappa\varDelta_F(x-x'), \tag{A·3}$$

$$\therefore \qquad (\Box-\kappa^2) \frac{\partial}{\partial\kappa} \frac{\partial}{\partial x_\lambda} \varDelta_F(x-x') = 2\kappa \frac{\partial}{\partial x_\lambda} \varDelta_F(x-x'),$$

$$\therefore \qquad \frac{\partial}{\partial\kappa} \frac{\partial}{\partial x_\lambda} \varDelta_F(x-x') = \frac{\kappa}{i} \int \varDelta_F(x-x'') \frac{\partial}{\partial x_\lambda''} \varDelta_F(x''-x') dx'',$$

$$\therefore \qquad \frac{\partial}{\partial\kappa} (\kappa^2 \varDelta_F(x-x')) = \frac{\kappa}{i} \int \frac{\partial}{\partial x_\lambda} \varDelta_F(x-x'') \frac{\partial}{\partial x_\lambda''} \varDelta_F(x''-x') dx''.$$

Furthermore

$$(\Box-\kappa^2)(x-x')_\lambda \varDelta_F(x-x') = 2 \frac{\partial}{\partial x_\lambda} \varDelta_F(x-x'),$$

$$\therefore \qquad (x-x')_\lambda \varDelta_F(x-x') = \frac{1}{i} \int \varDelta_F(x-x'') \frac{\partial}{\partial x_\lambda''} \varDelta_F(x''-x') dx'',$$

$$\therefore \qquad (x-x')_\lambda \frac{\partial}{\partial x_\lambda} \varDelta_F(x-x') = \frac{1}{i} \int \frac{\partial}{\partial x_\lambda} \varDelta_F(x-x'') \frac{\partial}{\partial x_\lambda''} \varDelta_F(x''-x') dx''$$
$$- 4\varDelta_F(x-x')$$

$$= \frac{1}{\kappa} \frac{\partial}{\partial\kappa} (\kappa^2 \varDelta_F(x-x')) - 4\varDelta_F(x-x'),$$

therefore

$$(x-x')_\lambda (x-x')_\lambda \varDelta_F(x-x') = \frac{1}{i} \int (x-x'')_\lambda \varDelta_F(x-x'') \frac{\partial}{\partial x_\lambda''} \varDelta_F(x''-x') dx''$$

$$+ \frac{1}{i} \int \varDelta_F(x-x'')(x''-x')_\lambda \frac{\partial}{\partial x_\lambda''} \varDelta_F(x''-x') dx''$$

$$= -\int \varDelta_F(x-x'') \frac{\partial}{\partial x_\lambda''} \varDelta_F(x''-x''') \frac{\partial}{\partial x_\lambda'''} \varDelta_F(x'''-x') dx'' dx'''$$

$$- \int \varDelta_F(x-x'') \frac{\partial}{\partial x_\lambda''} \varDelta_F(x''-x''') \frac{\partial}{\partial x_\lambda'''} \varDelta_F(x'''-x') dx'' dx'''$$

$$- \frac{4}{i} \int \varDelta_F(x-x'') \varDelta_F(x''-x') dx''$$

$$= \frac{1}{2\kappa^2}\left[-\frac{\partial^2}{\partial\kappa^2}(\kappa^2\varDelta_F(x-x')) - \frac{1}{\kappa}\frac{\partial}{\partial\kappa}(\kappa^2\varDelta_F(x-x')) - 4\kappa\frac{\partial}{\partial\kappa}\varDelta_F(x-x')\right]. (A\cdot4)$$

From (A·3), we obtain

$$(\Box-\kappa^2)\frac{\partial^2}{\partial\kappa^2}\varDelta_F(x-x') = 2\varDelta_F(x-x') + 4\kappa\frac{\partial}{\partial\kappa}\varDelta_F(x-x'),$$

$$\therefore \quad (\Box-\kappa^2)\frac{\partial^3}{\partial\kappa^2}(\kappa^2\varDelta_F(x-x')) = 2\kappa^2\varDelta_F(x-x') + 4i\partial(x-x') + 4\kappa\frac{\partial}{\partial\kappa}(\kappa^2\varDelta_F(x-x')),$$

On the other hand, we have

$$(\Box-\kappa^2)\frac{\partial}{\partial\kappa}(\kappa^2\varDelta_F(x-x')) = 2\kappa^3\varDelta_F(x-x') + 4i\kappa\partial(x-x'),$$

so

$$(\Box-\kappa^2)\frac{\partial}{\partial\kappa}(\kappa^2\varDelta_F(x-x')) - (\Box-\kappa^2)\left(9\kappa\frac{\partial}{\partial\kappa}\varDelta_F(x-x')\right)$$

$$= 4\kappa\frac{\partial}{\partial\kappa}(\kappa^2\varDelta_F(x-x')) - 16\kappa^2\varDelta_F(x-x') + 4i\partial(x-x')$$

and

$$4\kappa\left[\frac{\partial}{\partial\kappa}(\kappa^2\varDelta_F(x-x')) - 4\kappa\varDelta_F(x-x')\right]$$

$$= (\Box-\kappa^2)\left[\frac{\partial^2}{\partial\kappa^2}(\kappa^2\varDelta_F(x-x')) - 9\kappa\frac{\partial}{\partial\kappa}\varDelta_F(x-x') - 2\varDelta_F(x-x')\right]. \qquad (A\cdot5)$$

Since we have

$$(\Box-\kappa^2)\left[\frac{\partial}{\partial\kappa}(\kappa^2\varDelta_F(x-x')) - \kappa^2\frac{\partial}{\partial\kappa}\varDelta_F(x-x')\right] = 4i\kappa\partial(x-x'),$$

we obtain from (A·5) and this equation

$$\frac{\partial^2}{\partial\kappa^2}(\kappa^2\varDelta_F(x-x')) - 9\kappa\frac{\partial}{\partial\kappa}\varDelta_F(x-x') - 2\varDelta_F(x-x')$$

$$-\frac{1}{i}\int\left[\frac{\partial}{\partial\kappa}(\kappa^2\varDelta_F(x-x^2)) - \kappa^2\frac{\partial}{\partial\kappa}\varDelta_F(x-x^2)\right]$$

$$\times\left[\frac{\partial}{\partial\kappa}(\kappa^2\varDelta_F(x^2-x^1)) - 4\kappa\varDelta_F(x^2-x^1)\right]dx^2. \qquad (A\cdot6)$$

(2)   *On the equation* (A·3)

We generally obtain, from

$$(\Box-\kappa^2)\frac{\partial}{\partial\kappa}\varDelta_F(x-x') = 2\kappa\varDelta_F(x-x'),$$

$$\frac{\partial}{\partial\kappa}\varDelta_F(x-x') = a\varDelta^{(1)}(x-x') + b\varDelta(x-x') + \int\varDelta_F(x-x'-x'')\varDelta_F(x'')dx'',$$

but if

$$\Delta_F(x-x') = \begin{cases} -i\,\Delta^+(x-x') & \text{for } x_0 > x_0', \\ -i\,\Delta^-(x-x') & \text{for } x_0 < x_0' \end{cases}$$

is taken into account, have $a=b=0$.

The identity

$$\frac{\partial}{\partial \kappa}\Delta_F(x-x') = \frac{\kappa}{i}\int \Delta_F(x-x'')\Delta_F(x''-x')dx'',$$

which holds for $x \neq x'$, is also correct for $x=x'$, because from

$$\Delta_F(0) = \int \delta(x)\Delta_F(x)dx,$$

we obtain

$$\frac{\partial}{\partial \kappa}\Delta_F(0) = \int \frac{\partial}{\partial \kappa}\Delta_F(x)\delta(x)dx$$

$$= \frac{\kappa}{i}\int \Delta_F(x'')\Delta_F(x'')dx''.$$

References

1) P. A. M. Dirac, Proc. Camb. Phil. Soc. **30** (1934), 150.
  W. Heisenberg, ZS. Phys. **90** (1934), 209.
  R. Serber, Phys. Rev. **49** (1936), 545.
  V. Weisskopf, D. V. S. M. F. M., **14** (1936), 6.
  R. Peierls, Proc. Roy. Soc. **146** (1934), 420.
  Z. Koba, T. Tati and S Tomonaga, Prog. Theor. Phys. **2** (1947), 101, 198.
  J. Schwinger, Phys. Rev. **75** (1949), 651.
  G. Wentzel, Phys. Rev. **74** (1948), 1070.
  W. Pauli and F. Villars. Rev. Mod. Phys. **21** (1949), 434.
  J. Rayski, Acta Physica Polonica 9 (1948), 129.
  R. Jost and J. Rayski, Helv. Phys. Acta **XXII** (1949), 457.
  D. Feldman, Phys. Rev. **76** (1949), 1369.
  H. Umezawa and R. Kawabe, Prog. Theor. Phys. **4** (1949), 420, 423, 443.
  J. Schwinger, Phys. Rev. **82** (1951), 664.
  E. C. G. Stueckelberg and A. Petermann, Helv. Phys. Acta **XXVI** (1953), 499.
2) A Pais and S. T. Epstein, Rev. Mod. Rhys. **2** (1949), 445.
  K. Sawada, Prog. Theor. Phys. **5** (1950), 117.
  J. Yukawa and H. Umezawa, Prog. Theor. Phys. **6** (1951), 112.
  Y. Katayama, Prog. Theor. Phys. **5** (1950), 1054.
  F. Rohrlich, Phys. Rev. **77** (1950), 357.
  F. Villars, Phys. Rev. **79** (1950), 122.
  S. Borowitz and W. Kohn, Phys. Rev. **86** (1952), 985.
  Y. Takahashi and H. Umezawa, Prog. Theor. Phys. **8** (1952), 193.
3) H. Umezawa and R. Kawabe, Prog. Theor. Phys. **5** (1950), 769.
4) H. Fukuda and Y. Miyamoto, Prog. Theor. Phys. **4** (1949), 394.

Reduced Widths from (d, p) Reactions

Yoichi FUJIMOTO*

Department of Physics, Kyoto University

Ken KIKUCHI

Department of Physics, Osaka University

Shiro YOSHIDA

Research Institute for Fundamental Physics, Kyoto University

(Received December 30, 1953)

Reduced widths of low-lying levels are calculated from experimental deuteron stripping cross sections on the basis of Butler theory. The validity of the approximation is examined by these reduced widths. The structure of low-lying levels is discussed from our results.

§ 1. Introduction

The purpose of the present paper is to estimate the reduced widths of low-lying levels by analysing the experimental data of deuteron stripping reactions in reference to a number of available experiments, and to discuss the level structure on the basis of our results.

Butler's theory of deuteron stripping reactions has made a great success in determining quantum numbers of a captured nucleon and a final nuclear state. This theory is powerful also to obtain the knowledge of structures of a nuclear level. We can calculate the reduced widths of various nuclei from the experimental deuteron stripping cross section on the basis of Butler's theory, although the validity of the theory is necessary to examine. This reduced width is a very important quantity in particular when it is obtained from a (d, p) reaction. The reasons are as follows : (1) The reduced width of a low-lying level where nucleon scattering experiments are not applicable is obtained in a rather direct way. (2) The orbital angular momentum of a captured nucleon is known simultaneously. (3) When the level is of complicated structure which may have the two different orbital angular momenta, the reduced widths for different angular momenta can be obtained separately as pointed out by Bethe and Butler.[1]

In § 2 we summarize the procedure to calculate the reduced width and the validity of our approximations is discussed. In § 3 we list the reduced widths from available data

---

* Now at H. H. Wills Physical Laboratory, University of Bristol.

and structures of levels are discussed on the basis of our theory.

§ 2.   Derivation of a formula for cross sections and test of its approximation

In the previous paper[2] by one of the authors we obtained a formula which expresses the deuteron stripping cross section in reference to the Butler approximation. This formula can be obtained more easily by using the fact that the Butler approximation is equivalent to the modified Born approximation.[3] Here we shall present the procedure of calculation for instructive purpose.

First we consider the case in which the neutron in a deuteron is captured by a potential and all particles have zero spin. In the Born approximation the differential cross section is given by the well-known formula,

$$\frac{d\sigma}{d\Omega} = \frac{M^2}{2\pi^2 \hbar^4} \frac{k_p}{k_d} \sum_m |<f| V_n |i>|^2 \tag{1}$$

where M is the nucleon mass, k_d is the wave number of the incident deuteron, k_p the one of the outgoing proton and V_n the interaction potential. $|i>$ and $<f|$ are given by

$$|i> = e^{ik_d(r_n + r_p)/2} \varphi_d(r) \tag{2a}$$

$$<f| = e^{-ik_p r_p} Y_{lm}^*(\theta_n \varphi_n) 1/r_n R_l^*(r_n) \tag{2b}$$

where r_n and r_p are the coordinates of the neutron and the proton respectively, $r = r_n - r_p$, and $Y_{lm}(\theta_n \varphi_n) 1/r_n R_l(r_n)$ is the wave function of the captured neutron in the state of an angular momentum l and Z component m. If we use the deuteron wave function of the Hulthén type,

$$\varphi_d(r) = N_d(e^{-ar} - e^{-\beta r})/r \tag{3}$$

the matrix element $<f|V_n|i>$ is obtained as

$$<f|V_n|i> = 4\pi N_d \left(\frac{1}{a^2 + K^2} - \frac{1}{\beta^2 + K^2}\right) \delta_{m,0} v^l \sqrt{4\pi(2l+1)}$$

$$\times \frac{1}{Z} \int j_l(Zr_n) V_n(r_n) R_l(r_n) dr_n \tag{4}$$

where $K = 1/2k_d - k_p$, $Z = k_d - k_p$, and $j_l(x) = \sqrt{\pi x/2} J_{l+1/2}(x)$. The potential $V_n(r_n)$ can be eliminated by using the Schrödinger equation for the captured neutron and we get the following expression for the integral in equation (4),

$$-\frac{\hbar^2(x_n^2 + Z^2)}{2MZ} \int R_l(r_n) j_l(Zr_n) dr_n,$$

where $x_n = \sqrt{2M|E_n|}/\hbar^2$ and E_n is the binding energy of the neutron in final state. Here we restrict the region of integration over r_n only to the outer region according to the modified Born approximation. Then the above expression is reduced to

$$-\frac{\hbar^2}{2MZ}R_l(a)\left[\frac{\partial j_i(Zr_n)}{\partial r_n}-g_i j_i(Zr_n)\right]_{r_n=a},$$ (5)

a is the nuclear radius by which the external and the internal region are divided. g_l is the logarithmic derivative of $R_l(r_n)$ at $r_n=a$ and can be evaluated with the external wave function and the list of g_l for a neutral particle is given in reference 2. Thus we get the cross section in terms of $R_l(a)$. Here we remove the restrictions and consider a real nucleus. If the target nucleus is of spin j and the final nucleus of J, then the cross section must be multiplied by $(2J+1)/2(2j+1)(2l+1)$ and the mass M must be reduced mass and K and Z are given by

$$K=1/2k_d-k_p, \qquad Z=k_d-\frac{M_i}{M_i+M_n}k_p,$$ (6)

respectively, where M_i is the mass of a target nucleus and M_n is the one of a neutron. The wave function at the nuclear radius $a^{-1}R_l(a)$ must be replaced by the following integral

$$1/aR_l(a)\rightarrow[\int\Psi_f^*(r_n,x)\Psi_i(x)dx]_{r_n=a}$$ (7)

where $\Psi_i(x)$ is the product of the target nucleus wave function and the spin-angle wave function of neutron and Ψ_f is the final state wave function and $\int dx$ is the integral over all coordinates of final nucleus but r_n and the sum over spins.

On the other hand the reduced width of a nuclear level for channel c is defined as[4]

$$\gamma_c^2=\frac{\hbar^2 a^2}{2M_c}\left|\int[\Psi_f^*(r_n,x)\Psi_i(x)dx]_{r_n=a}\right|^2\Big/\int_{r\le a}|\Psi_f(r_n,x)|^2 r_n^2 dr_n dx$$ (8)

where M_c is the reduced mass for channel c. Thus $R_l(a)$ must be replaced by

$$|R_l(a)|^2\rightarrow\frac{2M_c}{\hbar^2}\gamma_c^2(1+\sum_c\gamma_c^2\dot{g}_c)^{-1},$$ (9)

where $\dot{g}_c$ is the energy derivative of the logarithmic derivative g_l and its expression for some l are given in reference 2. The final expression for (d,p) reactions is now given by

$$\frac{d\sigma}{d\Omega}=\frac{2J+1}{2j+1}\cdot\frac{k_p}{k_d}\cdot\frac{M_f}{M_f+M_n}\frac{M_i}{M_i+M_d}4\pi N_d^2$$

$$\times\sum\left|\frac{1}{Z}\left(\frac{1}{\alpha^2+K^2}-\frac{1}{\beta^2+K^2}\right)(j_i'-g_i j_i)\right|^2\frac{2M_c}{\hbar^2}\gamma_c^2(1+\sum_c\gamma_c^2\dot{g}_c)^{-1}$$ (10)

where M_d and M_f are the masses of deuteron and final nucleus respectively, and j_i' is the derivative of j_i with respect to r_n at the nuclear surface. The reduced width is often expressed in unit of the sum rule limit $(a/3)(2M/\hbar^2)\gamma^2$, denoted by θ^2.

Before going to the application of (10) we shall discuss the choice of parameters and the validity of our approximations. Our expression for the cross section involves the nuclear radius a, and the magnitude of a is chosen in such a way that the theoretical

angular distribution agrees best with the experimental one. In comparison with an experimental cross section we take the differential cross section at the first maximum. We believe that this way of comparison is least influenced by our approximation.

Since Butler's approximation neglects various effects, it is important to know to what extent our approximation is valid. Thomas[5] has tried to test the approximation by Butler on the basis of the charge symmetry. Since the reduced widths of corresponding levels of mirror nuclei should have the same value, we can compare these two reduced widths, one calculated from the (d, p) reaction and the other from (p, p) scattering. The results shown in Table 1 and 2 indicate that the two kinds of reduced width differ by factor 5.

Table 1. Comparison of the reduced widths of C^{13} and N^{13} from $(dp)^*$ and (pp) reaction. References are given Table 4.

Configuration	C^{13}		N^{13}	
	E_x (Mev)	Θ^2 from (dp) %	E_x (Mev)	Θ^2 from (pp) %
$2s_{1/2}$	3.099	12	2.369	54
$(2p_{3/2})^{-1}(2p_{1/2})^2$	3.68	0.64	3.511	3.1
$3d_{5/2}$	3.89	14	3.558	21

* Ajzenberg and Lauritsen; Rev. Mod. Phys. **24**, 321.

Table 2. Comparison of the reduced widths of O^{17} and F^{17} from (dp) and (pp) reaction. References are given in Table 4.

Configuration	O^{17}		$F^{17}*$	
	E_x (Mev)	Θ^2 from (dp) %	E_x (Mev)	Θ^2 from (pp) %
$2s_{1/2}$	0.875	15	0.536	33

* Teichmann and Wigner; Phys. Rev. **87** 123

A similar test of the approximation is possible on the basis of the charge independence. The reduced widths of two levels of $T=1$ with $T_z=1$ and $T_z=0$ are not equal but the one for $T_z=1$ is as large as the one for $T_z=0$ by factor two. The levels of B^{12} at 0.95 Mev formed by the reaction $B^{11}(d, p)B^{12}$ and the level of C^{12} at 16.10Mev formed by the reaction $B^{11}(p, p)B^{11}$ would be a corresponding level.

Fujimoto, Ichikawa and Yamaguchi[6] have pointed out that there exists an uncertainty for the absolute value of (d, p) cross sections, which comes from the ignorance of interaction between the proton in an incident deuteron and a target nucleus in the Butler approximation. Horowitz and Messiah[7] also showed independently that this discrepancy between the two corresponding reduced widths, one from (d, p) and one from (p, p) reactions, may be removed by considering the above mentioned proton-nucleus interaction and assuming this interaction as repulsive. Further they showed that the reduction of the cross section due to this repulsion is more important, the smaller the angular momentum of a captured neutron is.

Table 3.   Comparison of the reduced widths of Mg^{25} and Al^{25} from (d,f) and (d,n) reaction.

Configuration	Mg^{25}		Al^{25}*	
	E_x (Mev)	θ^2 from (df) %	E_x (Mev)	θ^2 from (dn) %
$3d_{5/2}$	0	3.9	0	0.28
$2s_{1/2}$	0.582	8	0.45	1.7
$3d_{3/2}$	0.976	3.4	0.95	0.29

* Goldberg; Phys. Rev. **89**, 760

Next we compare, the reduced widths from (d, p) and (d, n) reactions for the same target nucleus. In Table 3 the results for Mg^{24} (d, p) Mg^{25} and Mg^{24} (d, n) Al^{25} are shown. A pair of the reduced widths should have the same magnitude according to the charge symmetry but they actually differs by factor ten. Only their relative magnitudes in one reaction show a similar tendency to those of the other. It is difficult to understand such a large difference by the difference of the Coulomb barrier for (d, p) and (d, n), for the incident energy of deuterons for both cases are higher than the barrier height. We see that (d, p) and (d, n) reactions can not be treated equally. Therefore we shall not treat the reduced width from (d, n) reactions in the next section. The absolute value, as seen from the preceding discussions may not be correct by factor five or so, but the relative value for various nuclei may be more confidential. This can be suggested from the fact that the reduced widths from (d, p) reactions, in Table 1 are consistently smaller than those from (p, p) reactions, but their relative magnitudes for three levels are similar for both reactions.

§ 3.   Results and discussions

Following the procedure presented in the last section we calculated the reduced widths from available experimental data of (d, p) reactions and showed the results in Table 4. In column 1 the types of reactions, in column 2 final excitation energies, in column 3 the angular momenta of captured neutrons, in column 4 differential cross sections at the first maxima, in column 5 nuclear radii a adopted, in column 6 the spins of final states are given. The columns 8 and 9 show reduced widths and their value in unit of the sum rule limit, θ^2. θ^2 is a measure of validity of the single particle model as pointed out by Teichmann and Wigner.[1] When θ^2 is close to unity the configuration of the level is well approximated by a single particle picture, whereas the configuration is of complicated structure for θ^2 much smaller than unity. We shall discuss this point in individual cases.

First we are concerned with the case where a target nucleus is even-even. It is seen that θ^2 in some levels are larger than those of other levels in the same nucleus. The level of large θ^2 is such that a neutron is captured in a new orbit which is not yet occupied by the nucleons in the target nucleus and further the captured level lies energetically far from the last orbit of the target nucleus. For example C^{12} has filled shells $1s_{1/2}$ and $2p_{3/2}$. The orbits $2s_{1/2}$ and $3d_{5/2}$ are new ones and energetically far from $2p_{3/2}$ orbit.

Hence the 3.099 Mev and 3.89 Mev levels, that are supposed to be orbits $2s_{1/2}$ and $3d_{5/2}$ respectively, may have θ^2 larger than others. On the other hand 3.68 Mev level is of orbit $(2p_{3/2})^{-1}$, so that the orbit is already occupied by nucleons in the target nucleus. This will explain why θ^2 of the level is smaller than others by about factor ten. Similar examples are found in other even-even nuclei.

The neglected effects in the Butler approximation may however change even the relative value of θ^2. If the calculation by Horowitz and Messiah[7] is a good approximation, γ^2 in Table 4 should be multiplied by a factor which increases with decreasing angular momentum of a captured neutron l. Since the large θ^2's listed in Table 4 often appear for small l's, this tendency may be further amplified.

The behaviour of θ^2 derived from (d, p) reactions may be in similar situation to such a characteristic behaviour of ft values of β-decays that is explained by de-Shalit and Goldhaber[8] in terms of the purity of configuration. Nuclear states with the same spin and parity can be mixed due to configuration interactions. The order of mixing is supposed to be proportional to the matrix element between the states concerned and to the inverse of energy difference between these two states. The matrix element is large when the two states have analogous configurations. When the captured neutron has a new and energetically distant orbit the configuration is purified and a single particle model turns to be a good approximation. Since the reduced width, as seen from its definition (8), involves an overlapping integral of initial and final states, it should be larger, the better the overlap is. On the other hand when the mixing of states is so considerable that the level can not be expressed by those of a target nucleus and a captured neutron the overlapping becomes bad and θ^2 is reduced. 3.68 Mev level of C^{13} is an example of this, as investigated by Lane[9] who got a reasonable agreement of θ^2 with that predicted by intermediate coupling.

Next we turn to the case where a final nucleus is even-even. As Stähelin and Preiswerk[10], de-Shalit and Goldhaber[8] and other authors pointed out, the low-lying excited states of an even-even nucleus have the same single particle orbit as its ground state but of much different structures. The excited state has a seniority number different from that of the ground state and has a complicated structure in which the configuration of two excited neutrons is mixed with that of two excited protons. Therefore the reduced width of the excited state must be much different from that of the ground state. Such nuclei listed in Table 4 are too few to argue the general trend. It is desirable that more experimental data will be accumulated.

Now we shall study the variation of reduced widths from nucleus to nucleus. In Fig. 1 the variation of θ^2 with neutron number N is shown. The levels chosen for a nucleus are those which are considered to be the most purified ones. The peaks of θ^2 occur when N approaches the number at which a shell or a subshell is filled. This is similar to the behaviour of the excitation energies of the first excited states for even-even nuclei. As pointed out by Stähelin and Preiswerk[10] the number of states produced by nucleons in an unfilled shell is largest when nucleons fill about a half of shell or sub-shell. Then the possibility increases that two states with the same spin and parity lie close to each other so that the level has a mixed configuration. Consequently the reduced width is much

reduced and the level distance becomes small.   In terms of Bohr's collective model,[11] the deformability of a core is much smaller for a closed shell than that for an unclosed shell

Fig. 1   The variation of θ^2 with neutron number N

and the overlapping of initial and final states is good when one nucleon is added to a target nucleus.   It is noted that θ^2's of Ca^{41} is unexpectedly small in spite of the fact that the nucleus has the configuration of closed shell for both neutrons and protons plus a neutron and the single particle model should be a good approximation.   But it is compatible with the fact that the first excitation energy of even-even nuclei does not show a marked peak when the proton number approaches to 20 as Stähelin and Preiswerk[10] showed.

From a statistical point of view Feshbach, Peaslee and Weisskopf[12] noticed that there exists a relation between the reduced width for a certain channel of a nucleus and the distance of levels for that channel.   We tried to examine the relation, presenting the ratios of averaged reduced width to the level distance D in column 10 of Table 4.   These level distances, which are shown in column 9, are estimated by averaging the distances of several lowest excited states irrespective of their spin, parity and channel. γ^2/D is thought to be approximately proportional to the wave number of a neutron in a nucleus and to be of almost the same magnitude for various nucleus.   As seen from this table there exists discrepancy by up to a factor 50, and this difference is a little larger than in the case of nucleon scattering.[4]   This may be due to the fact that we treat the low-lying level for which the statistical fluctuation may be considerable.

Table 4. Reduced widths from (d,p) reaction.

Reaction	E_x (Mev)	l	$d\sigma/d\Omega(\theta)$ (mb)	$a(10^{-13}$ cm)	J	γ^2 $(10^{-13}$ Mev cm)	$\theta^2(\%)$	D (Mev)	γ^2/D 10^{-13}cm	Configuration	Reference
$_4\mathrm{Be}^9_5(d,p)$ $_4\mathrm{Be}_6^{10}$	0	1	3.1(30°)	4.5	0	0.76	5.5	3.0	0.15		a
	3.37	1	4.45(30°)		2	0.12	0.88				
$_6\mathrm{C}_6^{12}(d,p)$ $_6\mathrm{C}_7^{13}$	0	1	40(18°)	4.0	1/2	1.4	9 3	1.0	1.4	$2p_{1/2}$	b
	3.099	0	103(16°)		1/2	1.9	12			$2s_{1/2}$	
	3.68	1	5(30°)		3/2(1/2)	0.10	0.64(1.3)			$(2f_{3/2})^{-1}$ $(2p_{1/2})^2$	c
	3.89	2	76(30°)		5/2(3/2)	2.2	14(21)			$3d_{5/2}$	
$_7\mathrm{N}_7^{14}(d,p)$ $_7\mathrm{N}_8^{15}$	0	1	11.8(18°)	5.6	1/2	1.3	12	1.5	4.1	$(2p_{1/2})^{-1}$	
	7.16 7.31	0	150(22°)		1/2, 3/2	[16, 3.4]	[150,31]			$s_{1/2}$	d
	8.32	0	130(18°)		3/2(1/2)	7.2	65(250)			$s_{1/2}$	
	9.22	0	30(20°)		3/2(1/2)	8.6	77(540)			$s_{1/2}$	
$_8\mathrm{O}_8^{16}(d,p)$ $_8\mathrm{O}_9^{17}$	0	2	30(36°)	5.0	5/2	0.65	5.2	1.0	1.1	$3d_{5/2}$	e
	0.875	0	12.5(45°)	6.0	1/2	1.6	15			$2s_{1/2}$	
$_9\mathrm{F}_{10}^{19}(d,p)$ $_9\mathrm{F}_{11}^{20}$	0	2	0.92(45°)	5.32	(1)	0.077	0.66	0.26	1.5		f
		0	1.01(4°)			0.0049	0.042				
	0.65	2	7.15(45°)		1~3	0.50~0.22	4.3~1.9				
	2.05	2	<8(45°)		1~3	<1.5~0.22	<4.3~1.9				
	3.49 3.53	0	116(7°)		0,1	1.4, 0.36 [0.70, 0.18]	12, 3.1 [6,1.6]				
$_{12}\mathrm{M}_{12}^{24}(d,p)$ $_{12}\mathrm{Mg}_{13}^{25}$	0	(2)	17.0(38°)	5.3	2/5	0.46	3.9	0.5	1.1	$3d_{5/2}$	h
	0.582	0	230(0°)		1/2	0.96	8			$2s_{1/2}$	
	0.976	2	9.3(30°)		3/2(5/2)	0.40	3.4(2.3)			$3d_{3/2}$	
	1.957	2	8.9(32°)		3/2, 5/2	0.27, 0.18	2.3, 1.5				
	3.405	1	140(15)°		3/2, (1/2)	0.89	7.6(16)			$p_{3/2}$	
$_{12}\mathrm{Mg}_{13}^{25}(d,p)$ $_{12}\mathrm{Mg}_{14}^{26}$	0	2	1.6(30°)	5.3	0	2.2	19	0.8	2.0		h
	1.825	0	17(0°)		2,3	0.18 0.13	1.5, 1.2				
		2	11(35°)			2.7, 1.9	23, 16				
	2.972	0	160(0°)		2,3	1.8, 1.2	15, 10				
	3 969	0	170(0°)		2,3	1.8, 1.2	15, 10				
	4.353	0	190(0°)		2,3	1.9, 1.3	16, 11				
	6.147	0	120(0°)		2,3	0.75, 0.54	6.4, 4.6				
$_{11}\mathrm{Mg}_{14}^{26}(d,p)$ $_{12}\mathrm{Mg}_{15}^{27}$	0	0	110(0°)	5.3	1/2	0.43	3.7	1.1	0.32	$2s_{1/2}$	h
	0.887	2	8.7(30°)		3/2(5/2)	0.27	2.3(1.6)			$3d_{3/2}$	
$_{13}\mathrm{Al}_{14}^{27}(d,p)$ $_{13}\mathrm{Al}_{15}^{28}$	0(d)	0	220(0°)	6.13	3(2)	1.1 [0.55]	11(16) [5.5, 8 0]	0.3	5.9		g
	1.0	0	19(0°)		2,3	0.15, 0.10	1.5, 1.0				
		2	11(30°)			1.4~0.87	14, 8.6				
$_{14}\mathrm{Si}_{14}^{28}(d,p)$ $_{14}\mathrm{Si}_{15}^{29}$	0	0	62(0°)	5.4	1/2	0.25	2.2	0.7	0.25	$2s_{1/2}$	i
	1.278	2	6.2(32°)		3/2 (5/2)	0.23	2.0(1.3)			$3d_{3/2}$	
	2.027	2	2.4(32°)		5/2 (3/2)	0.082	0.71(1.1)			$(3d_{5/2})^{-1}$ $(2s_{1/2})^2$	
	3.070	2	1.2(30°)		3/2, 5/2	0.023, 0.015	0.20, 0.13				
	3.623	3	4(45°)		7/2 (5/2)	0.14	1.2, (1.6)			$4f_{7/2}$	
	4.934	1	55(12°)		3/2 (1/2)	0.36	3.1, (6.2)			$3p_{3/2}$	
	6.380	1	32(12°)		1/2 (3/2)	0.32	2.8, (1.4)			$3p_{1/2}$	

Reaction	E_x (Mev)	l	$d\sigma/d\Omega(\theta)$ (mb)	$a(10^{-13}$ cm)	J	$\gamma^2(10^{-13}$ Mev cm)	$\theta^2(\%)$	D (Mev)	γ^2/D 10^{-13}cm	Configuration	Reference
$_{18}A_{22}^{40}(d,p)$ $_{18}A_{23}^{41}$	0	3	2.7(36°)	5.87	7/2 (5/2)	0.10	0.9 (1.2)	0.5	0.19	$4f_{7/2}$	d
$_{20}Ca_{20}^{40}(d,p)$ $_{20}Ca_{21}^{41}$	0	3	3.8(40°)	5.87	7/2 (9/2)	0.15	1.4 (1.1)	0.7	0.22	$4f_{7/2}$	j
	1.90	1	23(18°)		3/2 (1/2)	0.22	2.0 (4.0)			$3p_{3/2}$	
	2.42	1	10(18°)		1/2 (3/2)	0.17	1.6 (0.8)			$3p_{1/2}$	
	3.96	2	6.0(20°)		3/2 (5/2)	0.18	1.7 (1.1)				
	4.76	2	7.2(23°)		3/2 (5/2)	0.13	1.2(0.80)				
	5.72	2	7.8(23°)		3/2 (5/2)	0.089	0.84(0.56)				

a　Fulbright et al., Phys. Rev. **88**, 700, $Ed=3.6$ Mev
b　Rotblat, Nature, **167**, 1027, $Ed=8$ Mev
c　Rotblat, Phys. Rev. **83**, 1271, $Ed=8$ Mev
d　Gibson et al., Proc. Roy. Soc. **210**, 543, $Ed=8$Mev
e　Burge et al., Proc. Roy. Soc. **210**, 534, $Ed=7.7$ Mev
f　Bromley et al, Phys. Rev., **89**, 396, $Ed=3.6$ Mev
g　Holt and Marsham, Proc. Phys. Soc. **66A**, 249, $Ed=8$ Mev
h　Holt and Marsham, Proc Phys. Soc. **66A**, 258 $Ed=8$ Mev
i　Holt and Marsham, Proc. Phys. Soc., **66A**, 467 $Ed=8$ Mev
j　Holt and Marsham, Proc. Phys. Soc., **66A**, 565, $Ed=8$ Mev
　　The Figures in square brackets are averages over doublet levels.

References.

1)　H. A. Bethe and S. T. Butler, Phys. Rev. **85** (1952), 1045.
2)　S. Yoshida, Prog. Theor. Phys. **10** (1953), 1, 370
3)　E. Gerjuoy, Phys. Rev. **91** (1953), 645.
4)　T. Teichmann and E. P. Wigner, Phys. Rev. **87** (1952), 123.
5)　R. G. Thomas, Phys. Rev. **91** (1953), 453.
6)　Y. Fujimoto, Y. Ichikawa and Y. Yamaguchi; Private communication.
7)　J. Horowitz and A. M. I. Messiah Report of Birmingham Conference p. 41 (1953) and Jour. de Phys. et Rad. **14** (1953), 693.
8)　A. de-Shalit and M. Goldhaber, Phys. Rev. **92** (1953), 1211.
9)　A. M. Lane, Report of Birmingham Conference (1953), Proc. Phys. Soc. **66A** (1953) 977.
10)　P. Stähelin and P. Preiswerk, Nuovo Cimento **10** (1953), 1219.
11)　A. Bohr and B. R. Mottelson, Dan. Mat. Fys. Medd. **27** (1953), No. 16.
12)　H. Feshbach, D. C. Peaslee and V. F. Weisskopf, Phys. Rev. **71** (1947), 145.

[Note added in proof; It has been kindly remarked by Dr. Lane that the same value of the nuclear radii should he used for (d,p) and (p,p) reactions at the same target nuclei, whereas the reduced widths in Table 1 and 2 were calculated with different radii for these two reactions. Therefore the test of the validity of the Butler theory is only of an approximate meaning, although the radius for (d,p) reactions does not strongly affect the reduced width.]

On the Structure of the Interaction of the Elementary Particles, V

—— *Interaction of the Second Kind and Non-local Interaction* ——

Susumu KAMEFUCHI

Institute of Theoretical Physics, Nagoya University

By extending the same idea as in the previous paper IV to the case of vertex corrections we have examined to what extent the theory can be made less singular with such a method. Detailed discussions are given as to the easiest example, namely, the 4-vertex-interaction in the meson pair theory. For this interaction it is required to introduce the infinite number of counter terms, which can, however, be brought into the form of the non-local interactions. Thus, although we can not say that the theory thus obtained is closed, we can obtain the finite results if we admit such a procedure. Lastly, some remarks are briefly stated in connection with other approaches to the non-singular theory.

§ 1. Introduction

It is a well-known fact that for the interaction of the second kind the closed theory can not be obtained as far as we resort to the usual perturbation method together with the renormalization idea. We have seen in the previous paper II) that the straightforward application to this case of the renormalization method leads to the introduction of infinitely many number of interactions (counter terms) with the high derivatives and the strong non-linearity, which are thus to be regarded as non-local interactions. To examine the behavior of such counter terms in order to study how to generalize this method so as to be applicable to interactions of the second kind, it is one of the purposes of the present paper. It has been reported by Stueckelberg and Petermann[2] that in such an attempt the counter terms can be obtained for the quantum electrodynamics of Proca particles, its details, however, being not yet published.

The first difficulty we encounter on carrying out this program is that it becomes necessary to introduce the infinite number of counter terms with the nature mentioned above, or in other words, as proceeding to the higher order approximation any type of graph becomes divergent. In the previous paper IV we have investigated a method to evade such a difficulty and to construct a closed theory for interactions of the second kind. In general, when an interaction of the second kind is present, the field reaction appearing in the higher order approximation not only leads to the strong resistive effect, but also to the correspondingly strong inductive one. In order to construct a less singular theory we have, thus, taken advantage of this fact and examined the prescription of using the corrected propagation function instead of the usual uncorrected one. We have found there that it is generally required to introduce the auxiliary interactions of the second kind for

the closure of the theory. We may suppose that such an idea will not be restricted only to the correction of the propagation function, but it may also be extended to the method of the correction of the vertex. To investigate to what extent such a method can be used to make the theory less singular, it is the second purpose of this paper. This method is convenient for our first purpose, too. Recently, Edwards[3] has adopted the similar idea to obtain the completely convergent theory for interactions of the first kind. When applied to the interaction of the second kind, such an approach is very effective in suppressing the divergency, that is, the divergence degree and non-linearity of diverging graphs.

In § 2 the difficulty of this method is pointed out when applied to the 3-vertex interaction. Its origin may be partly due to the use of the perturbation approach. It is further shown that for the 4-vertex interaction some kind of correction can lead to the corrected vertices having the cut-off denominator and in § 3 they are applied to the evaluation of the S-matrix. It is found there that we can not restrict the number of primitive divergences for the general interactions of this type, excepting that of the meson-pair theory. In § 4 the detailed discussion is given concerning the meson-pair theory to give the concrete forms of non-local counter terms. In the last section § 5 the concluding remarks and the further outlook are stated in connection with the approach to our problems.

§ 2.   The vertex correction

In this section we shall consider the 3- and 4-vertex interactions of the second kind, by which we mean, as usual, the interaction Hamiltonian having three and four field operators, respectively.

First we shall consider the 3-vertex interaction. For the same reason as in the previous paper, in order to obtain the approximate corrected vertex part with the enough cut-off factor to make the theory less singular, we have to select a set of special kind of graphs generated by the iteration of an appropriate graph and express their total effect in a condensed form. More mathematically speaking, this is equivalent to solve an appropriate integral equation for the amplitude corresponding to the vertex part. When carrying through our program for a given 3-vertex interaction we must consider the higher order graphs belonging to some categories of vertex part, which are not reducible to the line corrections, that is, the graphs as shown in Fig. 1. In this paper we use the words

Fig. 1

"category" and "categorization" in the following meanings: We say that when some graphs for vertex part are made of a vertex part plus a graph generated by iterating a given graph some times, they belong to the same categorization and that if a graph is the one corresponding to the n-time iteration, it is a graph of category n.

But, as far as we are obliged to treat the problem by the perturbation method, we usually encounter the following several difficulties. First, there arises the troublesome situation concerning the so-called overlapping divergences, which is also the case for

interactions of the first kind. Thus, for example, when constructing a higher order graph (belonging to a higher category) by connecting several convergent constituents, there appear new types of divergences. Such divergences always depend not only on the part of the graph concerned, but also the construction of the whole graph, which require to be treated from the overall view-point. Therefore, in order to extract the finite parts with some significances it would be necessary to introduce new kinds of counter terms. Of course, as mentioned above, such a circumstance arises for interactions of the first kind in so far as we use the perturbational method*. (But, there the further introduction of new kind of counter terms is unnecessary, because the order of divergence is never raised as proceeding to the higher order categories of graphs, and so all the divergences are managed merely by adjusting the coupling constants of counter terms already given in the lower categories.) In the case of interactions of the second kind in which we are interested, we need to modify the operator form of counter terms, and so it becomes difficult to collect the total finite remainders and express them in a concrete and compact form.

Another difficulty comes from the situation peculiar to the present case. When combining two constituent graphs of some category, both of which are already made finite, to construct a graph belonging to a higher category, further divergences anew arise from the integrations with respect to the connecting internal lines. This circumstance superimposed on the first one leads to the perplexing complications. For these reasons, when starting with the 3-vertex interaction as the primary one, we may say that it is very difficult to write the corrected vertex parts and the corresponding counter terms in a concrete form as in the case of the previous paper, and so we can not obtain the closed non-singular theory by use of our manipulation.

Next, we shall apply the analogous consideration to the 4-vertex interaction. As examples of such interactions we know the Fermi-interaction of β-decay and the pair inter-action in the meson theory, etc. If we adopt the same categorization as above the answer will also be given in the negative. But, there is another possibility, looking the other way.

For the 4-vertex interaction we now consider the vertex parts belonging to the categorization as shown in Fig. 2a. In this case the overlapping divergent integrals do not appear and so, if each constituent of the lower category, for instance, as shown in Fig. 2b be convergent after the subtraction, the final expression will also be convergent. Hence, in order to make finite all the graphs belonging to our categorization by means of counter terms, it seems at first sight only sufficient to introduce a single counter term which subtracts the divergence corresponding to Fig. 2b. But, the following consideration shows that the situation is not so simple for counter terms, while it is quite evident that the whole finite remainder can easily be obtained in a simple form. As to this point we shall give
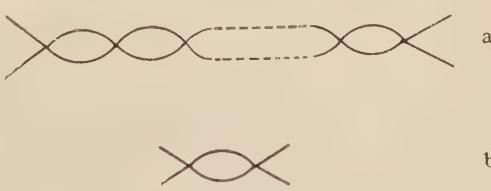
Fig. 2

* Recently, Edwards has made a non-perturbational approach to the similar problems. See reference 3.

a more detailed consideration in the following.

As the primary interaction Hamiltonian we shall assume the form

$$H(x) = f F_1(\partial) \phi^1(x) F_2(\partial) \phi^2(x) F_3(\partial) \phi^3(x) F_4(\partial) \phi^4(x) + \text{c.c.}, \tag{1}$$

where F_α's $(\alpha = 1, 2, 3, 4)$ are functions of the differentiation operators acting on $\phi^\alpha(x)$. Here, we further add an assumption that $\phi^\alpha \neq \phi^3 (a \neq 3)$, which somewhat simplifies the discussion. Let a^α be the highest degree of the derivation operators in $F_\alpha(\partial)$, i.e.,

$$F_\alpha(\partial) \sim \partial^{a^\alpha} \quad \text{for large momenta.} \tag{2}$$

The characteristic constants for the interaction $(1)^{1)}$ are defined respectively by

$$
\left.
\begin{aligned}
A &= \sum_{\alpha=1}^{4} a^\alpha, \\
B &= \sum_{\alpha=1}^{4} b^\alpha, \\
C &= 4
\end{aligned}
\right\} \tag{3}
$$

Hence, the dimension of the coupling constant f is given by

$$\eta = A + \tfrac{1}{2} B. \tag{4}$$

At first, we shall take the graph of category 1 as shown in Fig. 3a, in which we assume the contracted internal lines belong to the fields 3 and 4 (or 1 and 2), respectively and the four external lines to the fields 1 and 2 (or 3 and 4), respectively. For definiteness of the discussion we further prescribe that the vertex correction is always made in such a way that the contractions are performed with respect to the fields 3 and 4 (or 1 and 2) in pairs. Thus, for instance, such a graph as contracted with respect to fields 1 and 3 is not considered here, which is taken into account in the second stage of the procedure, i.e., in the stage where we calculate the S-matrix in terms of corrected vertices.

Fig. 3

Now, let us return to our problem. The amplitude [1] corresponding to the graph of category 1 is written in the form

$$[1] = F_1(\partial) \phi^1 F_2(\partial) \phi^2 II(34) F_1(\partial) \phi^1 F_2(\partial) \phi^2. \tag{5}$$

As is easily seen, the degree of divergence of the matrix element II (34) is given by $(2a^3 + 2a^4 + b^3 + b^4)$ (or the value with the replacement $3, 4 \to 1, 2$). II is further rewritten in the form

$$II(34) = D(3, 4) + C(34), \tag{6)*}$$

where D and C are divergent and convergent parts, respectively, of the left member. Thus,

---

* There is much arbitrariness in this separation of divergent part. In this respect, see III.
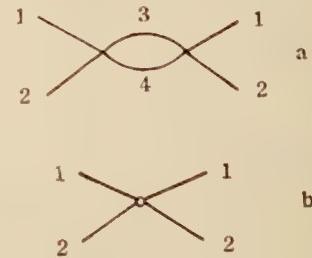
corresponding to the separation (6), the required counter term is to be of the form

$$-F_1\phi^1 F_2\phi^2 D(34) F_1\phi^1 F_2\phi^2, \tag{7}$$

which we represent as the graph as shown in Fig. 3b and regard as the graph of category 1. On considering both graphs 3a and 3b we obtain the finite part

$$F_1\phi^1 F_2\phi^2 C(34) F_1\phi^1 F_2\phi^2. \tag{8}$$

For the vertex part with external lines 3 and 4 the discussions are made along the similar line. In what follows, the matrix element of the graph with four external lines referring to fields α, β, γ and δ, respectively, will be called the $(\alpha\beta)(\gamma\delta)$ amplitude.

Fig. 4

Next, we proceed to the category 2, the graphs as shown in Fig. 4a, whose external lines belong to 1, 2, 3 and 4, respectively, and so, which can be regarded as the radiative correction to the primary interaction (1). The corresponding matrix element is written as
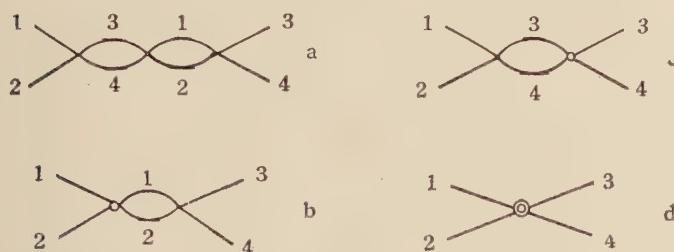
$$F_1\phi^1 F_2\phi^2 [\mathit{\Pi}(34)\mathit{\Pi}(12) - D(34)\mathit{\Pi}(12) - \mathit{\Pi}(34)D(12)] F_3\phi^3 F_4\phi^4. \tag{9}$$

The second and the third terms come from the graphs 4b and 4c, respectively, involving the counter terms to remove divergences arising from the subintegrations $\mathit{\Pi}(34)$ and $\mathit{\Pi}(12)$, respectively. The sum of these three graphs, however, is not convergent in itself. This is due to the fact that the counter terms, having external lines, lead to divergences when connected with other graphs. In this point lies, we think, one of the most characteristic differences between the methods of renormalization and of regularization. For in the latter each part of an expression can be treated separately and the final expression may automatically become finite provided that all parts of the graph are convergent. In order to remove the remaining divergent term from the expression (9) we further need the following counter term

$$F_1\phi^1 F_2\phi^2 D(34)D(12) F_3\phi^3 F_4\phi^4. \tag{10}$$

When proceeding in this way to the higher categories (Fig. 2a) and summing up the contributions coming from the graphs formed in all possible ways of combination of the primary interaction and those counter terms, which have been introduced in the lower categories, we see that it is required to introduce a new counter term for each category. In this respect it appears that the situation here is rather similar to the case of 3-vertex, but it differs from the latter in the following points : (1) that the finite parts can collectively be brought into a condensed expression and (2) that this is also the case for counter terms, since those of high category can always be expressed in terms of the ones of lower categories. In connection with our first requirement for the closed theory,

however, we must say that the closed theory can not be obtained, since the number of necessary counter terms is, in fact, infinite.

The final expressions for the finite remainders, including up to the infinite time iterations can be written in the forms

$$F_1\phi^1 F_2\phi^2 \frac{C(34)}{1-C(12)C(34)} F_1\phi^1 F_2\phi^2 , \quad \text{etc.,} \tag{11}$$

$$F_1\phi^1 F_2\phi^2 \frac{1}{1-C(12)C(34)} F_3\phi^3 F_4\phi^4. \tag{12}$$

In what follows, we shall term these two corrected vertices $(12)(12)$ vertex and $(12)(34)$ vertex, respectively. The required counter terms as a whole can be brought into the form

$$F_1\phi^1 F_2\phi^2 \frac{-D(34)}{1-D(12)D(34)} F_1\phi^1 F_2\phi^2, \quad \text{etc.,} \tag{13}*)$$

$$F_1\phi^1 F_2\phi^2 \frac{D(34)D(12)}{1-D(34)D(12)} F_3\phi^3 F_4\phi^4. \tag{14}$$

The general form of the counter terms as given by (13) and (14) can easily be inferred from the similar consideration as mentioned above but with slight extensions to the general case. Here, we give in this connection a more convenient method, since we shall meet, in § 4, a more general case, to which this method is to be applied.

Now, let us consider a general graph belonging to the category n, remembering that such a graph has no overlapping divergences. On separating its matrix element $[n]$ into the divergent and convergent parts by applying the Salam's method[1] we can write it as follows ;

$$[n] = D(\overbrace{C+D)\cdots(C+D}^{(n-1)}) + (C+D)D(\overbrace{C+D)\cdots(C+D}^{(n-2)}) + \cdots + (\overbrace{(C+D)\cdots(C+D}^{(n-1)})D$$

$$- DD(\overbrace{C+D)\cdots(C+D}^{(n-2)}) - \cdots - D(C+D)D(\overbrace{C+D)\cdots(C+D}^{(n-3)}) -$$

$$+ DDD(\overbrace{C+D)\cdots(C+D}^{(n-3)}) + \cdots \tag{15}$$

$$+ \cdots$$

$$+ (-1)^{n+1}\overbrace{DD\cdots D}^{n} + \text{convergent part,}$$

where the arguments (12), (34) in the factors have been omitted, but they shall be understood without any confusion, since all the factors are written down in the same order as in the corresponding graph (Fig. 2). The sign attached to each divergent factor is

* These expressions are to be symbolically understood, since D involves the derivation operators. The same remarks apply to the expressions (33)—(36), below.

due to the general property of the so-called isolated divergence. All divergent terms take the forms such as (divergent constant) × (reduced integral). From the eq. (15) we see that the negative of each diverging term makes one-to-one correspondence with the contribution from a graph in which the factors such as $DD\cdots D$ are all replaced by the corresponding counter terms (for example, Fig. 4b, c, d). In other words, for instance the diverging term $D\cdots D$ can be cancelled by the counter term $-F\phi F\phi D\cdots DF\phi F\phi$ and generally the term $D\cdots(C+D)\cdots D\cdots$ can be compensated by a graph involving in it several counter terms. Thus, in order to cancel all the divergent terms we need the counter terms with the form $(-1)^i F\phi F\phi \overbrace{DD\cdots D}^{i} F\phi F\phi$ $(i=1, 2,\cdots, n)$. Among these counter terms those for $i < n$ all refer to the divergences of the sub-integrations in $[n]$ and so they are nothing but the terms to be introduced in the lower categories $i(i < n)$. Therefore, when proceeding in this way from the lower categories to the higher step by step we have only to introduce, for the category n, a single counter term $(-1)^n F\phi F\phi DD\cdots DF\phi F\phi$, which corresponds just to the divergence of the final integration. That is, according as we take a new graph, it is only sufficient to introduce a new counter term which removes the divergence from the final integration. From these observations we can easily find the final expression (13) and (14). This method can also be applied to the case involving the overlapping divergences (§ 4).

§ 3. The divergences in the S-matrix

As seen in the previous section, in an attempt to obtain a closed non-singular theory by use of the corrected vertices we need the infinite number of counter terms even in the stage of constructing the corrected vertices. Now, making a further step, at any rate, on allowing such a case that the necessary counter terms can be obtained in a compact form, there arises then a question as to whether it is possible or not to obtain a non-singular theory under such conditions.

Setting aside, for a moment, the problem whether all the elements of S-matrix can in fact be rewritten in terms of the corrected vertices, to which we shall return in the next section, we shall first examine to what extent the theory can be made less singular when we use the corrected vertices. For this purpose, we have to obtain the characteristic constants referring to the modified interactions (11) and (12) (the constants corresponding to (3) and (4)).

In the corrected vertex part we shall denote by p_μ the sum of 4-momenta referring to two external lines in pairs of the fields 1 and 2 (or 3 and 4). As is easily seen, the asymptotic behavior of the (12)(12) amplitude (8) in high energy region takes the form $C(12) \sim p^{2(a^1+a^2)+b^1+b^2}$ and so we find

$$A_{(12)(12)}=2(a^1+a^2)-2(a^{1'}+a^{2'})-(b^1+b^2), \tag{16}$$

whence

$$\eta_{(12)(12)}=0. \tag{17}$$

Similarly,

$$\eta_{(34)(34)}=0, \tag{18}$$

while for the $(12)(34)$ vertex (12) we have

$$A_{(12)(34)} = \sum_{\alpha=1}^{4} a^{\alpha} - \{2\sum_{\alpha=1}^{4} a^{\alpha} + \sum_{\alpha=1}^{4} b^{\alpha}\}, \tag{19}$$

and so

$$\eta_{(12)(34)} = A_{(12)(34)} + \frac{1}{2}\sum_{\alpha=1}^{4} b^{\alpha} = -\eta. \tag{20}$$

At first sight from (17), (18) and (20) it appears that since all modified vertices assume the behavior of the interaction of the first kind we can easily obtain the non-singular theory. But, we must notice the following fact : As the cut-off factor in the denominators of (11) and (12) is a function with respect to p_{μ}, it does not play the role of the high-momentum cut-off if two external lines ϕ_{1} and ϕ_{2}, say, in pairs are incident on a same vertex, for the denominator then becomes a function of external momenta and not of virtual ones. Therefore, the analysis of the divergence becomes somewhat complicated in this case.

For convenience in the following discussion we shall call " the special vertex " such a vetex that does not cut-off the high momenta. The special vertex can be placed between two graphs or more generally several graphs can be connected in series by such vertices. For such a graph we shall cut it into several segments at the points of intermediary special vertices and consider each segment separately. As for the whole graph itself connected by special vertices we shall return to it later on.

Concerning a given graph, let n be the number of $(12)(34)$ vertices, where the cut-off denominator is effectively operating and let n_{12} and n_{34} be numbers of special vertices on which external lines ϕ_{1}, ϕ_{2} and ϕ_{3}, ϕ_{4} are incident in pairs. Then, the condition of the primitive divergence is written as follows :

$$\sum_{\alpha=1}^{4}\left(a^{\alpha} + \frac{b^{\alpha}}{2} + 1\right)E^{\alpha} + n\left(\sum_{\alpha=1}^{4} a^{\alpha} + \frac{1}{2}\sum_{\alpha=1}^{4} b^{\alpha}\right) - (n_{12} + n_{34})\left(\sum_{\alpha=1}^{4} a^{\alpha} + \frac{1}{2}\sum_{\alpha=1}^{4} b^{\alpha}\right) \leq 4. \tag{21}$$

Now, we shall put

$$\left. \begin{aligned} E^{1\text{ or }2} &= E^{1\text{ or }2\prime} + n_{12}, \\ E^{3\text{ or }4} &= E^{3\text{ or }4\prime} + n_{34}, \end{aligned} \right\} \tag{22}$$

where $E^{\alpha\prime}$ is the number of external lines of α-field incident on the ordinary vertices other than the special ones. Inserting (22) in (21) we find

$$\sum_{\alpha=1}^{4}\left(a^{\alpha} + \frac{b^{\alpha}}{2} + 1\right)E^{\alpha\prime} + n\left(\sum_{\alpha=1}^{4} a^{\alpha} + \frac{1}{2}\sum_{\alpha=1}^{4} b^{\alpha}\right)$$

$$+ n_{12}\left(2 - a^{3} - a^{4} - \frac{b^{3}+b^{4}}{2}\right) + n_{34}\left(2 - a^{1} - a^{2} - \frac{b^{1}+b^{2}}{2}\right) \leq 4. \tag{23}$$

From this result we immediately see the following facts. If the constants a^{α} and b^{α} take the sufficiently large values, all the graphs with $n_{12} = n_{34} = 0$ shall be convergent. (Of

course, in this respect one must remember that we are assuming that all the vertices can be replaced by the corrected ones.) But, we shall notice that among the graphs with non-vanishing n_{12} and n_{34} but with vanishing n, there exist those graphs having arbitrarily large values of E^x's. In order to limit the number of diverging graphs of this kind we must impose the following conditions:

$$\left. \begin{aligned} F_{12} &\equiv 2 - a^1 - a^2 - \frac{b^1 + b^2}{2} \geq 0, \\ F_{34} &\equiv 2 - a^3 - a^4 - \frac{b^3 + b^4}{2} \geq 0. \end{aligned} \right\} \tag{24}$$

The interactions of the second kind allowed by these conditions are as follows:

$$\left. \begin{aligned} &g(\bar{\psi}^1 \Omega \partial_\mu \psi^2)(\bar{\psi}^3 \Omega \partial_\mu \psi^4), \quad &&\eta = 4, \; F_{12} = F_{34} = 0; \\ &g(\bar{\psi}^1 \Omega \partial_\mu \psi^2)(\bar{\psi}^3 \Omega \gamma_\mu \psi^4), \quad &&\eta = 3, \; F_{12} = 0, \; F_{34} = 1; \\ &g(\bar{\psi}^1 \Omega \psi^2)(\bar{\psi}^3 \Omega \psi^4), \quad &&\eta = 2, \; F_{12} = F_{34} = 1; \end{aligned} \right\} \tag{25}$$

<div align="center">etc.</div>

$$\left. \begin{aligned} &\lambda \partial_\mu \phi^1 \partial_\mu \phi^2 \partial_\nu \phi^3 \partial_\nu \phi^4, \quad &&\eta = 4, \; F_{12} = F_{34} = 0; \\ &\lambda \partial_\mu \phi^1 \cdot \phi^2 \cdot \partial_\mu \phi^3 \cdot \phi^4, \quad &&\eta = 2, \; F_{12} = F_{34} = 1; \\ &\lambda \partial_\mu \phi^1 \partial_\mu \phi^2 \cdot \phi^3 \phi^4, \quad &&\eta = 2, \; F_{12} = 0, \; F_{34} = 2; \end{aligned} \right\} \tag{26}$$

<div align="center">etc.</div>

$$\left. \begin{aligned} &f \bar{\psi}^1 \partial_\mu \psi^3 \partial_\mu \phi^2 \cdot \phi^4, \quad &&\eta = 2, \; F_{12} = F_{34} = \tfrac{1}{2}; \\ &f \bar{\psi}^1 \gamma_\mu \psi^3 \partial_\mu \phi^2 \cdot \phi^4, \quad &&\eta = 2, \; F_{12} = F_{34} = \tfrac{3}{2}; \\ &f \bar{\psi}^1 \psi^3 \phi^2 \phi^4, \quad &&\eta = 1, \; F_{12} = F_{34} = 3/2; \\ &f \bar{\psi}^1 \psi^2 \phi^3 \phi^4, \quad &&\eta = 1, \; F_{12} = 1, \; F_{34} = 2; \end{aligned} \right\} \tag{27}$$

<div align="center">etc.</div>

On the other hand, we must remember that in the above discussion those graphs have been cut and considered separately, which are connected by the special vertices. For such graphs there may occur a case in which, for example, two segments directly combined by an intermediary special vertex are both divergent. Under such circumstances the whole expression would not be made finite by use of two counter terms which compensate respective divergences coming from two segments. There usually remains the divergence arising from the final integration. This is just for the same reason as in § 2, where it was required to introduce new counter term for each category.

From the interactions, for which holds the relation

$$2(F_{12} + F_{34}) \leq 4, \tag{28}$$

we get the diverging graphs which have external lines more than 4. When we combine

these diverging graphs each other by the special vertices in a complex way[*], the resulting graph then would again diverge. Thus, even if we start with the interactions which satisfy the condition (24), we shall obtain, if they, in addition, satisfy (28), many diverging graphs with unlimited number of external lines. Accordingly, we shall then have to introduce the infinite number of counter terms with unlimited degree of non-linearity. In these cases it becomes very complicated, though it may be not necessarily impossible, to obtain the actual forms of the counter terms. Therefore, we shall here impose on the interaction a further condition

$$2(F_{12}+F_{34}) > 4. \tag{29}$$

As is easily seen, only the last two interactions in (27), i.e., the interaction in the meson pair theory, can satisfy this condition, about which we shall give the detailed consideration in the next section.

§ 4. The meson pair theory

Since the third and the fourth interactions in (27) are essentially equivalent, we shall consider, in what follows, the former for convenience. Then, the condition (23), corresponding to the present case, can be written in the form

$$\tfrac{3}{2}E'_{sp}+E'_{b}+n+\tfrac{3}{2}(n_{12}+n_{34}) \leq 4. \tag{30}[**]$$

Noticing that the numbers of the external spinor or bose lines are always even, respectively and putting $n^{*}=n_{12}+n_{34}$, we can classify the diverging graphs that satisfy the above condition as follows :

$$n^{*}=0 \begin{cases} \text{(a)} \begin{pmatrix} E'_{sp}=2, \\ E'_{b}=0, \end{pmatrix} & \text{(b)} \begin{pmatrix} E'_{b}=2, \\ E'_{sp}=0, \end{pmatrix} \\ \text{(c)} \begin{pmatrix} E'_{b}=4, \\ E'_{sp}=0, \ n=0 \ ; \ \text{Fig. 5a,} \end{pmatrix} \end{cases}$$

$$n^{*}=1 \quad \text{(d)} \begin{pmatrix} E'_{sp}=1, \\ E'_{b}=1, \ n=0 \ ; \ \text{Fig. 5b,} \end{pmatrix}$$

$$n^{*}=2 \quad \text{(e)} \begin{pmatrix} E'_{sp}=0, \\ E'_{b}=0 \ ; \ \text{Fig. 5c.} \end{pmatrix}$$
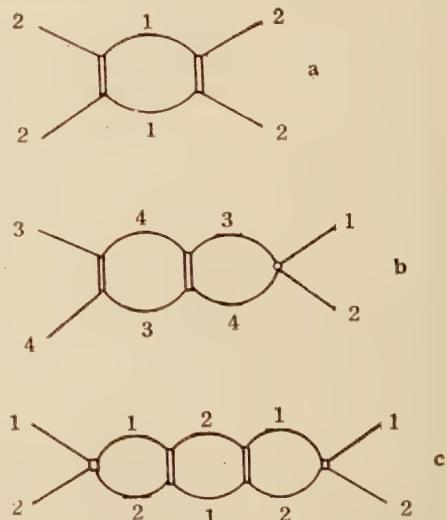
[] : (12)(12) vertex, (34)(34) vertex ;
□ : special vertex

Fig. 5

* This remark refers only to the connection between pairs of lines ϕ_1, ϕ_2 or ϕ_3, ϕ_4. For the connection between pairs of lines ϕ_1, ϕ_3 ; ϕ_2, ϕ_4. etc. there arises no trouble, since then the effective cut-off factor gives the convergent expression. Thus, we have only to take care of those graphs, which we have cut-off in the analysis of divergence.

** The graphs, which satify this condition, reducible or irreducible, lead to the divergences in the final integration.

So far we have assumed that the elements of the S-matrix can all be described in terms of the corrected vertices. We find, however, that there exists a graph for which this assumption is not legitimate. This exception is the lowest order self-energy graph of the field α, which has three internal lines. In such a graph we have not to correct both vertices, but only the one vertex, otherwise there is involved more than the terms we require. This exceptional case, however, does not give rise to any difficulty, because we can prove that such a self-energy term results in the corrected propagation functions, S_F' and $\varDelta_F'$ which have the same momentum dependence as the uncorrected ones, that is, $S_F'(p) \sim p^{-1}$, $\varDelta_F'(p) \sim p^{-2}$. Hence, how many self-energy graphs of this kind may be inserted in a graph, the integral thus obtained will have the same asymptotic momentum dependence as the integral before the insertion, and so as far as the divergence is concerned, there arises no problem under these conditions.

Further, it is a remarkable fact that in this type of self-energy graph there appear two virtual momenta t_1 and t_2, but contrary to the usual case with uncorrected vertices we can choose the variables in such a way that t_1 and t_2 sub-integrations converge and only the t_1 t_2 integration gives divergences, for the subtraction of which it is only sufficient to use the local mass type counter terms[*].

After all, the divergences we have to manage are the terms (a),$\cdots$, (e) as mentioned above, which we shall consider one by one. (a) and (b) correspond to the mass and the charge renormalization (Z_3-factor) terms and so can be treated by the usual prescription. (c) describes the boson-boson scattering process, being removed by the counter term such as $i\delta^2\phi^2\phi^2\phi^2$ familiar in the meson theory. Since this divergence is the logarithmic one and the attached external lines are boson lines, it does not give rise to any new difficult situation when combined with other diverging graphs. The graphs (d) and (e), however, requires much consideration. The counter terms to remove these divergences are of the same forms as those introduced when we construct the corrected vertices. If we connect these graphs mutually or through the intermediary of other corrected vertices, then we may have a case that the resulting graph is again divergent as a whole. Thus, we must subtract it once more. These circumstances are again due to the fact that in the previous section we considered the segments separately, the combination of which gives as a result divergent graph. Since these divergent terms are all of the form of 4-vertex amplitude, we shall consider, as the example, the (12)(12) amplitude.

Though it is an orthodox way to connect the divergent graph by use of the corrected vertices, we find it convenient to disentangle the original corrected vertices and then reconsider them term by term together with newly appearing divergent terms. That is, as far as the mutual connection of diverging graphs in series is concerned, we use the primary uncorrected vertices along with the corresponding counter terms. It should be noticed, however, that such a detour is merely a mathematical trick.

---

* For the uncorrected self-energy graph, it is required to introduce either the 4-vertex counter terms (referring to sub-integrations) and the mass type terms (referring to the final integration) or the non-local mass term.

In the general $(12)(12)$ graph and, in particular, in the term (e) the so-called overlapping divergences appear at the places of corrected vertices. To avoid this difficulty, we cut the $(12)(12)$ graph at the places of connecting ordinary vertices and separate into several elementary units in the way that there remain no longer any overlapping divergences between these elementary units. Then, we find that the elementary units lying at the end-points of the entire graph are either the term (d) or (e) and the intermediate units are the term (e). Since the property of the term (d) is easily seen from that of the term (e), we shall examine the latter in more details.

Among the divergent graphs belonging to (e), according to the number of connecting $(34)(34)$ vertices, we shall call the graph that of "category n" and denote its amplitude by $\{n\}$. As mentioned above, $\{n\}$ involves the overlapping divergences, and so, following Salam[4] we separate the divergent part in the form

$$
\begin{aligned}
\{n\} = &\mathfrak{D}_1^l R_{n-1} + \mathfrak{D}_2^l R_{n-2} + \cdots + \mathfrak{D}_n^l R_0 \\
&+ R_{n-1}\mathfrak{D}_1^r + R_{n-2}\mathfrak{D}_2^r + \cdots + R_0 \mathfrak{D}_n^r \\
&- \mathfrak{D}_1^l R_{n-2}\mathfrak{D}_1^r - \mathfrak{D}_2^l R_{n-3}\mathfrak{D}_1^r - \cdots \\
&+ \cdots \\
&+ \mathfrak{D}_{n\ final} + \text{convergent part,}
\end{aligned} \tag{31}
$$

where R_i is the reduced integral, $\mathfrak{D}_i$ the logarithmically divergent constant and the attached suffices refer to the order of category. In this equation the factors in each term are all placed in the order of the corresponding graph. Hence, the diverging terms are compensated by the corresponding graphs containing appropriate counter terms in just the same way as in § 2. A slight difference here is only seen in the fact that, for instance, $\mathfrak{D}_2^r \neq \mathfrak{D}_1^r \mathfrak{D}_1^r$ and so $\mathfrak{D}_2$ becomes an essentially new constant. For the later use we shall define the following constants :

$$
\left.
\begin{aligned}
\mathfrak{D}^r &= \mathfrak{D}_1^r + \mathfrak{D}_2^r + \cdots , \\
\mathfrak{D}^l &= \mathfrak{D}_1^l + \mathfrak{D}_2^l + \cdots , \\
\mathfrak{D}_{final} &= \mathfrak{D}_{1 final} + \mathfrak{D}_{2 final} + \cdots .
\end{aligned}
\right\} \tag{32}
$$

Now, we shall return to the problem of the general $(12)(12)$ graph which is formed by connecting several graphs (d) and (e). As remarked in § 2, here we remember that for a new graph concerned we have only to introduce the counter term corresponding to the divergence of the final integration and that the counter term which involves i factors of D or $\mathfrak{D}$ is to be attached the sign $(-1)^i$. In general, the $(12)(12)$ graphs are classified into four kinds according to the kinds of vertices lying at both ends of graph, that is ; (A) both ends are the ordinary vertices ; (B) ((C)) the left (right) end is the ordinary vertex and the right (left) one $(12)(12)$ vertex ; (D) both ends are (12) (12) vertices. Corresponding to these graphs, we obtain, noticing the above remarks, the following counter terms :

for (A) $\bar{\zeta}'^1\phi^2\left[\dfrac{-(D(34)+\mathfrak{D}_{fin}(34))}{1-(D(12)+\mathfrak{D}_{fin}(12))(D(34)+\mathfrak{D}_{fin}(34))}\right]\zeta'^1\phi^2.$ (33)

for (B) $\bar{\zeta}'^1\phi^2\left[-1+\dfrac{(D(34)+\mathfrak{D}_{fin}(34))}{1-(D(12)+\mathfrak{D}_{fin}(12))(D(34)+\mathfrak{D}_{fin}(34))}\right]\mathfrak{D}'\zeta'^1\phi^2,$ (34)

for (C) $\bar{\zeta}'^1\phi^2\mathfrak{D}'\left[-1+\dfrac{(D(34)+\mathfrak{D}_{fin}(34))}{1-(D(12)+\mathfrak{D}_{fin}(12))(D(34)+\mathfrak{D}_{fin}(34))}\right]\zeta'^1\phi^2,$ (35)

for (D) $\bar{\zeta}'^1\phi^2\mathfrak{D}'\left[\dfrac{-(D(34)+\mathfrak{D}_{fin}(34))}{1-(D(12)+\mathfrak{D}_{fin}(12))(D(34)+\mathfrak{D}_{fin}(34))}\right]\mathfrak{D}'\zeta'^1\phi^2.$ (36)

For the amplitudes such as $(12)(34)$ we can also obtain the counter terms in quite the same way.

The finite remainder of the amplitude $(12)(12)$ etc. can further be inserted in the other graph, but as seen from (30) this does not lead to any new circumstances as to the divergency of the latter graph. Besides such corrections, we have the self-energy ones. Therefore, in the general case the constants D and $\mathfrak{D}$ must be replaced with the corrected ones D^* and $\mathfrak{D}^*$, respectively. After all, under our prescription, the required counter terms are summarized in the self-energy terms, boson-boson scattering term and the terms such as $(33), \cdots, (36)$. We may say that the latter interactions are of the non-local form, since the constant D is a linear function of the momentum p.

Lastly, we must add a mathematical remark concerning the poles of the corrected vertices. In (11) and (12) it may occur that the denominator $A(p)=1-C(12)C(34)$ vanishes for some value of p. Let the mass of the field α be κ^α and assume for simplicity that $\kappa^1=\kappa^3$, $\kappa^2=\kappa^4$. If $p^2 \leq -(\kappa^1+\kappa^2)^2$, then $A(p)$ will have the imaginary part and so $A(p)$ does not in general vanish in this region, while for p^2 such as $p^2 > -(\kappa^1+\kappa^2)^2$ the imaginary part does not appear. On the other hand, the above mass relation gives the relation $C(12)=C(34)$, from which follows $A=1-(C(12))^2$. $C(12)$ can be regarded as a linear function with respect to the finite part λ_c's of the coupling constant $(\lambda_c+\partial\lambda)$ of the counter terms removing the divergence from II. Thus, there exists a possibility that by adjusting these parameters λ_c's we can take $|c|>1$, that is, we can avoid the poles by taking advantage of the arbitrariness in the counter terms. Under these conditions we can then easily show by a rather simple argument that the denominator A is sufficient enough to secure the convergence to such an extent as required in the foregoing sections.

§ 5. Concluding remarks

As an attempt to obtain the closed non-singular theory for the interaction of the second kind, the similar idea as developed in IV has been applied in this paper to the method of vertex corrections. That is, by modifying the usual renormalization theory in following that we first obtain the corrected vertices and then, as far as circumstances permit, we calculate the elements of S-matrix in terms of these vertices, we have examined the problem to what extent the theory can be made less singular. The corrected vertices,

having sufficient cut-off factor, have been obtained here as a result of iteration of certain matrix element, which we have obtained by the perturbation method. Under these circumstances it was found that the desired results can only be obtained for some special interaction, i.e., that of the meson pair theory, while for other interactions the situation becomes very complicated. Moreover, even for the former case it was required to introduce the infinitely many number of counter terms, being, however, brought into a form of the non-local interactions. From the view-point, therefore, as stated in I we can not say that it is a closed theory in the sense that we must allow the infinite types of subtractions. When we admit, however, such a kind of subtraction we can show that the finite results can be in fact derived.

When considering the theoretical and experimental successes which the renormalization theory has achieved in the field of quantum electrodynamics, it might be supposed that the appropriate formal extension of this method may become one of the shortest cut to approach to the theory of interaction of the second kind. As shown in I and in the above, such an extension of the renormalization method to this kind of interactions necessarily requires the introduction of the non-local interaction as the primary ones. In this connection we may say that the theory of non-local interaction, being recently discussed by many authors[F], aims at the same object by following the opposite way of approach, or in other words, that if we reverse the logical structure of our method, then we may obtain the latter theory. One of the weak points in the non-local theory lies, we think, in the fact that the principle is quite unknown, which determines the so-called form factors. Hence, in this sense we expect from the view-point of the correspondence principle that our approach might give an important clue to determine the form factors. Thus, in this respect it is no exaggeration to say that the interaction of the second kind and the non-local interaction constitute, as it were, two different aspects of a single logical structure. This will also be expected from the fact that since the coupling constant of the former is of the dimension of some power of length, the elements are already involved in it, which correspond to the non-local extension in the latter theory. Thus, as to our example we can start with the interactions (11) and (12) by treating the problem as that of the non-local theory.

The first difficult situation we encounter in following such a program lies in that the actual form of the non-local interactions obtained in this way depends on the details of the procedure we employed. Such an arbitrariness will further be enlarged if we remember that the method as developed in IV will also have a possibility of giving the closed non-singular theory. How to overcome such a situation, it is one of the problems, which remain to be investigated.

The arbitrariness comes from the circumstance that according to the method of correction the divergency in the theory shows the diverse properties. We have pointed out in IV that the last fact may have an effect upon the convergence of the perturbation series. If it is true, we shall have a condition which restricts the details of our prescriptions.

In the first paper I of this series we have classified the interaction into the first and the second kinds from the view-point of the applicability of renormalization theory and

there found the essential difference between them.   In IV and the present paper, however, we have further seen that such a classification may, to some extent, depend on the approximation method we have used.   This point has also been noticed recently by Edwards[2].

On the other hand, as being repeatedly emphasized by Sakata we have to recognize that the theoretical advance is, in itself, of the structure of many different levels.   Mutually contradicting elements existing in the theory at the original level shall be unified in the theory at the higher level, so that the qualitative contrasts revealed by the classification made in the former theory may then be replaced by something quantitative in the latter level.   Thus, the significance of the classification made in the theory at a level will be found in its playing the role of clarifying where truly the problems lies to be solved and how to attack them in raising the theory onto the higher level.   As for our present case, therefore, if the theory, being capable of treating the interaction of the second kind can be established through the investigation of this kind of interaction, then it may also be a more rational theory for the interaction of the first kind.

Therefore, at the present status it may be the most desirable way to rewrite the theory, as far as possible, in such a way as to be immediately applicable to both kinds of interactions.   Recent investigation made by Edwards[3] is very interesting in this respect, but he has not yet succeeded in treating the interaction of the second kind.

References

1) S. Sakata, H. Umezawa and S. Kamefuchi, Prog. Theor. Phys. **7** (1952), 377. H. Umezawa, ibid. **7** (1952), 551 ; S. Kamefuchi and H. Umezawa, ibid. **8**(1952) 579 ; **9** (1953), 529. In the text these papers are referred to as I, II, III and IV, respectively, and the same notations and definitions are used sometimes without special references.

2) E.C.G. Stueckelberg and A. Petermann, Helv. Phys. Acta **24** (1951), 317.

3) S. F. Edwards, Phys. Rev. **90** (1953), 284.

4) A. Salam, Phys. Rev. **82** (1951), 217 ; **84** (1951), 416.

5) P. Kristensen and C. Møller, Kgl. Danske Videnskab. Selskab, Mat.-fys. Medd **27** (1952), No. 7. C. Bloch, ibid. **27** (1952), No. 8.

On an Expression for the Total Cross Section[*]

L. I. Schiff[**]

Stanford University, Stanford, California, U.S.A.

The well-known expression for the total cross section in terms of the imaginary part of the coherent forward scattered amplitude is derived in a simple and general way from wave theory. The derivation also shows that there is a "shadow remnant" beyond the actual shadow, within which the forward coherent scattering intensity is diminished, of approximate radius $(r/k)^{1/2}$, where r is the distance from the scattering center and $k = 2\pi/\lambda$ is the wave number. A qualitative understanding of the shadow remnant is obtained in terms of the uncertainty principle.

It is well known that the total cross section (elastic plus inelastic and absorption, or coherent plus incoherent) in any scattering process is proportional to the imaginary part of the elastic or coherent scattered amplitude in the forward direction.[1] The physical reason for this proportionality is that any reduction in the incident beam, which represents a loss by scattering or absorption, must be produced by interference between the incident plane wave and the coherent scattered wave in the forward direction, and hence must be a linear function of the forward scattered elastic amplitude. This suggests that the theorem referred to above can be established by a straightforward consideration of the incident and scattered waves. It is shown below that this can be done, and that the region in the forward direction over which substantial destructive interference occurs can easily be found. The relation of this interference region to shadow formation is also discussed. Finally, the situation is discussed qualitatively from the point of view of the uncertainty principle. All of the calculations and results are applicable to the relativistic case.

We write the total stationary wave function u as the sum of a coherent part u_c and an incoherent part w. Asymptotically, u_c consists of an incident plane wave and an outgoing scattered wave ;

$$u_c \sim e^{ikr\cos\theta} + f(\theta, \phi)\,\frac{e^{ikr}}{r}, \tag{1}$$

where r, θ, ϕ are spherical coordinates referred to the incident direction as polar axis, and $k = 2\pi/\lambda$ is the wave number. The coherent radial flux in the outward direction does not involve w, and is proportional to :

$$-i\bar{u}_c\frac{\partial u_c}{\partial r}+\text{c.c.}\sim 2k\{\cos\theta+r^{-2}|f(\theta,\phi)|^2$$

$$+\frac{1}{2}r^{-1}(1+\cos\theta)[f(\theta,\phi)e^{ikr(1-\cos\theta)}+\text{c.c.}]\},\qquad(2)$$

where "c.c." denotes the complex conjugate. Here terms of the form $r^{-2}\exp(ikr)$, which arise both from the f term in (1) and from neglected higher order terms in the asymptotic series, are dropped because they average to zero over a wave-length in the radial direction. In (2), the first term on the right side corresponds to the component of the plane wave flux in the radial direction, the second term to the usual elastic scattered flux, and the third term to the interference between the incident and scattered amplitudes. The first term shows that (2) must be multiplied through by $r^2/2k$ in order to get a differential cross section per unit solid angle. Then integration of the right side over the sphere yields :

$$\int|f(\theta,\phi)|^2d\Omega+\frac{1}{2}r[\int(1+\cos\theta)f(\theta,\phi)e^{ikr(1-\cos\theta)}d\Omega+\text{c.c.}].\qquad(3)$$

The first term of (3) is the total elastic cross section σ_{11}. The whole of (3) is the total outward flux expressed as a cross section : since there is conservation, the sum of (3) and the total inelastic cross section σ_{inel} (which includes absorption) must be zero. Thus :

$$\sigma_{\text{total}}=\sigma_{\text{el}}+\sigma_{\text{inel}}$$

$$=-\frac{1}{2}r[\int\int(1+x)fe^{ikr(1-x)}dx\,d\phi+\text{c.c.}],\qquad(4)$$

where $x\equiv\cos\theta$. We integrate the first integral on the right side of (4) by parts with respect to x, to obtain :

$$-\frac{1}{2}r\int_0^{2\pi}\int_{-1}^{1}(1+x)fe^{ikr(1-x)}dx\,d\phi=-\int_0^{2\pi}d\phi\left\{\frac{i(1+x)}{2k}fe^{ikr(1-x)}\Big|_{-1}^{1}\right.$$

$$\left.-\int_{-1}^{1}\frac{i}{2k}e^{ikr(1-x)}\frac{\partial}{\partial x}[(1+x)f]dx\right\}.\qquad(5)$$

The second term on the right side is of order $r^{-1}\exp(ikr)$, as can be seen by another partial integration, and so can be neglected. The first term is $2\pi f(0)/ik$, where $f(0)/r$ is the forward scattered amplitude, and hence is independent of ϕ. We thus obtain from (4) and (5) :

$$\sigma_{\text{total}}=\frac{2\pi}{ik}[f(0)-\bar{f}(0)]=\frac{4\pi}{k}Im[f(0)],\qquad(6)$$

where "Im" denotes the imaginary part. Equation (6) is the well-known expression for the total cross section.

Inspection of the integral in (4) shows that the main contribution comes from those values of x for which $kr(1-x)\lesssim 1$; for large r, this corresponds to values of θ less than or of order $\theta_s\equiv(kr)^{-1/2}$. Only within the angle θ_s with the forward direction do

the plane and scattered waves have nearly enough a fixed phase relationship to interfere destructively and reduce the incident flux, which is proportional to $\cos\theta$. The radius of the region over which this interference occurs is of order $a_s \equiv r\theta_s = (r/k)^{1/2}$.

It is reasonable to suppose that this destructive interference is related to the shadow cast by the scattering or absorbing object in the short wave length case, $ka \gg 1$, where a is the radius of the object. In this case there is a strong forward diffraction maximum with angular width of order $\theta_a = 1/ka$, within which the scattering amplitude is roughly constant and equal to $f(0)/r$. The shadow can only be "block" out to such a distance that it is filled in by diffraction from the edge of the object; this distance is of order $R \equiv a/\theta_a$, so that $R = ka^2$. We now note that at this distance, $a_s = (R/k)^{1/2} = a$, so that the interference region has about the same size as the object. Further, the interfering or imaginary part of the forward scattered amplitude is, from (6):

$$Im[f(0)]/R = k\sigma_{\text{total}}/4\pi R ,$$

which is of order $ka^2/R = 1$, since σ_{total} is of order a^2. We see then that in the neighbourhood of the extreme limit of the shadow, the region of destructive interference and the object have about the same size, and the scattered amplitude and the incident plane wave amplitude have about the same magnitude, so that this interference is responsible for the shadow.

For $r > R$, there is no shadow; but the interference region, of size $a_s > a$, may be regarded as the remnant of the shadow. This shadow remnant has an angular width θ_s which is less than θ_a, so that it appears as a region of diminished intensity in the center of the forward diffraction maximum. For $r < R$, the variation of $f(\theta, \phi)$ with angle is of importance, and the situation is too complicated to analyze in this simple way.

When we are dealing with particles, the existence of the shadow remnant, of angular size $\theta_s = (kr)^{-1/2}$ and linear dimensions $a_s = (r/k)^{1/2}$, may also be understood in terms of the uncertainty principle. The shadow remnant is the region within which the incident plane wave intensity is diminished by such an amount as to make up for the particles that are scattered or absorbed by the obstacle. Thus if we add particles to the incident beam at the plane of the obstacle in such a way that the wave packet which describes them has minimum transverse dimensions at a distance r in front of the obstacle, this packet should just fill in the shadow remnant and hence be of size a_s. We ask for a minimum packet since the incident wave has maximum coherence (infinite plane wave), and no observations are made over the distance r. If the packet has a transverse size δx at the obstacle, it has a transverse momentum uncertainty $\delta p_x \sim \hbar/\delta x$. The spread in the packet over the distance r is the product of r and the angular divergence $\delta p_x/p$, where $p = \hbar k$. Thus the final transverse size is of order

$$\varDelta x \sim \delta x + r\frac{\delta p_x}{p} \sim \delta x + \frac{r}{k\delta x} .$$

Variation of δx showe that $\varDelta x$ is a minimum when δx, and also $\varDelta x$, is of order $(r/k)^{1/2}$. This is just equal to a_s, the size of the shadow remnant.

The writer is indebted to Professor E. Teller for an interesting discussion of this paper.

References

1) See M. Lax, Phys. Rev. **78** (1950), 306, and other papers cited there; R. J. Glauber and V. Schomaker. Phys. Rev. **89** (1953), 667.

On the Construction of S-matrix in Lagrangian Formalism

Tsutomu IMAMURA, Sigenobu SUNAKAWA and Ryôyû UTIYAMA

Department of Physics, Osaka University

(Received December 31, 1953)

Some differential equations with respect to a coupling constant g are written down in Lagrangian formalism. By integrating these equations S-matrix can be derived in two different forms; one is expressed in terms of the chronological symbol T^* (or P^*); the other in terms of T (or P). This way of consideration is applied to the system of local fields with non-local interactions. According to the result thus obtained, it seems necessary to make a further investigation on a new condition which ensures the existence of S-matrix.

Introduction and summary

Recently Yukawa[1] proposed a theory of non-local fields as an attack to the solution of both the problem concerning the intrinsic structure of elementary particles and that of difficulty of divergences. We shall make a study in the present paper only on the second problem, i. e., the problem of divergences, since his theory is so much complicated and there remain several problems to be investigated before we arrive at the goals mentioned above. Fierz[2] and Yukawa himself showed that Yukawa's non-local field is equivalent to a mixture of several local fields interacting non-locally with each other, every one of which has its own mass and spin. From this point of view, the difficulty of divergences might be removed, if possible, in Yukawa's theory by means of both the mixed fields and the non-local interactions. Although it may be incorrect to deal with these two problems separately, we shall tentatively restrict ourselves only to the non-local interaction of local fields for the sake of simplicity.

In order to verify that the system is free from divergences, it seems to be important to give the explicit expression of S-matrix, but in the non-local case it is not sure whether S-matrix exist or not. Therefore, as a first step we must verify the existence of S-matrix and then construct it explicitly. It is customarily believed that we can not set up a state vector on an arbitrary space-like surface in the system with non-local interaction. According to investigations by many authors[3], however, one may be allowed to assume the hypothesis of adiabatic switch-on and -off in the remote past and future, provided that the form factor has the causal character from the macroscopic standpoint. Accordingly in these places non-local character of the field quantities disappears, and state vectors may be defined. In such a case, there should be an S-matrix which connects these two state-vectors mentioned above. (The condition that the form factor should belong to the so-called " normal class[4] " seems to be also necessary.) In constructing the S-matrix, however, it

may be impossible to pile up the effects of interactions by solving the Tomonaga-Schwinger equation as in the usual case, according to our point of view. Therefore, it is desirable to find a new equation from which S-matrix can be derived.

Schwinger's "New Theory of Quantized Field[5]" suggests us the new equation required. Since his theory is constructed within the frame-work of the local field, it is important to investigate how many parts we can separate from his postulates, being dependent on the local concepts, in other words, how far we can extend this theory in the non-local case. In this paper, an outline of his theory is given in § 1. In § 2 we examine some particular equations in the local theory with derivative couplings, and write them down in Lagrangian formalism, one of which is the new equation for S-matrix. (In what follows this equation is called (A).) When we start from these equations, we can get explicit expression for S-matrix in § 3, in which the T^* symbol is naturally derived. In § 4, we take two assumptions, i) the S-matrix exists in the non-local case, being defined by

$$\varphi^{out}(x) = S^{-1}\varphi^{in}(x)S \tag{B}$$

where $\varphi^{in}(x)$, $\varphi^{out}(x)$ are incoming and outgoing operators respectively, ii) the new equation (A) also holds in this case. Unfortunately, however, we get the conclusion that two assumptions (A) and (B) can not be compatible in case of a general form factor and, therefore, at least one of them must be incorrect. Bloch[6] pointed out that, if one wants to get the S-matrix which satisfies (B) in a form of a power series in the coupling constant g, the S-matrix can not be obtained in the 4th order of g. Therefore, the postulate (A) turns out to be meaningless in that case. Detailed investigation concerning this compatibility is given in the appendix II. In conclusion we believe that there must be a new severe condition for the form factor in order to make S-matrix exist in the non-local case.

§ 1 Out-line of Schwinger's theory

In the present paper the system of natural units, i. e., $\hbar = c = 1$ is used, and the coordinate of any world point is denoted by $x^\mu = (x, y, z, t)$ the tensor indices are raised and lowered in the usual manner (where $g^{00} = g_{00} = -1$). In what follows all the arguments are referred to an arbitrary space-like hyper-plane σ, the unit normal of which is denoted by n_μ (in a special system of coordinates, n_μ has components $(0, 0, 0, 1)$).

Now the out-line of Schwinger's therory is given as follows;

(i) *Existence of transformation function*: *There exists a function* $\langle \varphi(\vec{x}, \tau_1) = \varphi'(\vec{x})$, $\tau_1 \mid \varphi(\vec{x}, \tau_2) = \varphi''(\vec{x}), \tau_2 \rangle$ *which transforms a state* $\mid \varphi_2 = \varphi'', \tau_2 \rangle$, *in which an operator* $\varphi(\vec{x}, \tau_2) \equiv \varphi_2$ *has a distribution of eigenvalues* $\varphi''(\vec{x}) \equiv \varphi''$ *at a time* $\tau_2 (\tau = n_\mu x^\nu)$, *into a state* $\mid \varphi_1 = \varphi', \tau_1 \rangle$ *in which an operator* $\varphi(\vec{x}, \tau_1)$ *has a distribution of eigenvalues* $\varphi'(\vec{x}) \equiv \varphi'$ *at a time* τ_1. *Especially in case of* $\tau = \tau_1 = \tau_2$ *it is defined by*

$$\langle \varphi = \varphi', \tau \mid \varphi = \varphi'', \tau \rangle = \delta [\varphi' - \varphi''], \tag{1.1}$$

where $\delta[\varphi' - \varphi'']$ is a δ-functional.

(ii) *Matrix elements*: *An operator* $\varphi(x)$ *is defined by its matrix elements*

$$\langle \varphi = \varphi', \tau_1 \mid \varphi(x) \mid \varphi = \varphi'', \tau_2 \rangle,$$

where x *is an arbitrary world point.* In particular when $\tau = \tau_1 = \tau_2$, their values are given by

$$\langle \varphi = \varphi', \tau \mid \varphi(x) \mid \varphi = \varphi'', \tau \rangle_{x \text{ on } \tau} = \varphi'(x)\delta[\varphi' - \varphi''] \qquad (1\cdot 2)$$

(iii) *Combination law*: *Transformation functions and matrix elements obey the combination law of matrices*;

$$\langle \varphi_1 = \varphi', \tau_1 \mid AB \mid \varphi_3 = \varphi''', \tau_3 \rangle$$
$$= \int \langle \varphi_1 = \varphi', \tau_1 \mid A \mid \varphi_2 = \varphi'', \tau_2 \rangle d\varphi'' \langle \varphi_2 = \varphi'', \tau_2 \mid B \mid \varphi_3 = \varphi''', \tau_3 \rangle$$

(iv) *Schwinger's Principle*: We consider a classical action integral

$$I[\varphi_H(x)] = \int_{\tau_2}^{\tau_1} d\tau \int_{-\infty}^{\infty} d\sigma \, L[\varphi''(x), \varphi_\mu^H(x)], \qquad (1\cdot 3)$$

where we have assigned to every world point x in the domain of integration Ω an arbitrary c-number $\varphi''(x)$ in such a way as being continuously differentiable with respect to x. Thus I is a functional of $\varphi^H(x)$, the domain of integration and the form of Lagrangian. Let us consider following three infinitesimal variations :

a)
$$L \rightarrow \bar{L} = L + \Delta L.$$

With the variation of Lagrangian, an operator $\varphi(x)$ belonging to the system with L will vary to an operator $\bar{\varphi}(x)$ belonging to that with $\bar{L}$;

$$\varphi \rightarrow \bar{\varphi}(x) = \varphi(x) + \Delta_L\varphi(x).$$

b) The domain of integration is changed from (τ_2, τ_1) to $(\tau_2 + \Delta\tau_2, \tau_1 + \Delta\tau_1)$.

c) The value $\varphi''(x)$ assigned to each point is changed to $\varphi'' + \Delta^*\varphi''$;

$$\varphi''(x) \rightarrow \varphi''(x) + \Delta^*\varphi''(x).$$

As a result of a), b) and c), we get

$$\bar{\varphi}(\vec{x}, \tau + \Delta\tau) = \varphi''(\vec{x}, \tau + \Delta\tau) + \Delta^*\varphi''(\vec{x}, \tau + \Delta\tau)$$
$$\approx \varphi''(\vec{x}, \tau) + \Delta\varphi''(\vec{x}, \tau), \qquad (1\cdot 4)$$

where

$$\Delta\varphi''(x) = \Delta^*\varphi''(x) + \frac{\partial \varphi''(x)}{\partial \tau} \Delta\tau. \qquad (1\cdot 5)$$

Putting

$$\varphi''(\vec{x}, \tau_1) = \varphi'(\vec{x}), \quad \Delta\varphi''(\vec{x}, \tau_1) = \Delta\varphi'(\vec{x})$$
$$\varphi''(\vec{x}, \tau_2) = \varphi''(\vec{x}) \quad \text{and} \quad \Delta\varphi''(\vec{x}, \tau_2) = \Delta\varphi''(\vec{x}),$$

we have

$$\bar{\varphi}(\vec{x}, \tau_1 + \Delta\tau_1) = \varphi'(\vec{x}) + \Delta\varphi'(\vec{x}) \quad \text{on} \quad \tau_1 + \Delta\tau_1,$$

and

$$\bar{\varphi}(\vec{x}, \tau_2 + \Delta\tau_2) = \varphi''(\vec{x}) + \Delta\varphi''(\vec{x}) \quad \text{on} \quad \tau_2 + \Delta\tau_2.$$

Then, Schwinger's Principle requires the following equation ;

$$\Delta\langle\varphi_1 = \varphi', \tau_1 \mid \varphi_2 = \varphi'', \tau_2\rangle = i\langle\varphi_1 = \varphi', \tau_1 \mid \Delta I \mid \varphi_2 = \varphi'', \tau_2\rangle. \tag{I}$$

The left-hand side of this equation is defined by

$$\Delta\langle\varphi_1 = \varphi', \tau_1 \mid \varphi_2 = \varphi'', \tau_2\rangle$$

$$\equiv \langle\bar{\varphi}(\vec{x}, \tau_1 + \Delta\tau_1) = \varphi' + \Delta\varphi', \tau_1 + \Delta\tau_1 \mid \bar{\varphi}(\vec{x}, \tau_2 + \Delta\tau_2) = \varphi'' + \Delta\varphi'', \tau_2 + \Delta\tau_2\rangle$$

$$- \langle\varphi(\vec{x}, \tau_1) = \varphi', \tau_1 \mid \varphi(\vec{x}, \tau_2) = \varphi'', \tau_2\rangle,$$

whereas ΔI by

$$\Delta I \equiv \int_{\tau_2 + \Delta\tau_2}^{\tau_1 + \Delta\tau_1} d\tau \int_{-\infty}^{\infty} d\sigma \, \bar{L}[\bar{\varphi}^H] - \int_{\tau_2}^{\tau_1} d\tau \int_{-\infty}^{\infty} d\sigma \, L[\varphi^H]$$

$$= \int_{\tau_2}^{\tau_1} d\tau \int_{-\infty}^{\infty} d\sigma \, \{\Delta L[\varphi^H] + [L]_\varphi \Delta^*\varphi^H\} + F(\tau_1) - F(\tau_2),$$

where the following abbreviation have been used ;

$[L]_\varphi =$ Euler's expression,

$$F(\tau_1) = \int_{\tau_1 = \text{const}} d\sigma_1 (\pi\Delta\varphi^H - n_\mu T^\mu{}_\nu \Delta x^\nu),$$

$$\pi = n_\mu \partial L/\partial\varphi_\mu = \partial L/\partial\dot{\varphi}, \quad \dot{\varphi} = d\varphi/d\tau = -n^\mu \partial\varphi/\partial x^\mu, \quad \Delta x^\mu = -n^\mu \Delta\tau$$

$T^\mu{}_\nu =$ energy-impuls tensor density.

As is well known, one can derive q-number Euler-equations, Schrödinger equation and the canonical commutation relations ([C. R.]) by taking particular variations. Here it must be noted that, as in the usual theory, every field quantity is expressed by a functional of the canonical set of variables φ and π, the latter being defined by

$$\langle\varphi = \varphi', \tau \mid \pi(\vec{x}, \tau) \mid \varphi = \varphi'', \tau\rangle = \frac{1}{i} \frac{\partial}{\partial\varphi'} \partial[\varphi' - \varphi'']. \tag{1.6}$$

Accordingly, $\dot{\varphi}$ can be defined as a functional of π and φ by help of the equation of fields rather than as an independent field variable.

The important equations which interest us in the following sections are

$$\Delta_L\langle\varphi_1 = \varphi', \tau_1 \mid \varphi_2 = \varphi'', \tau_2\rangle = i\langle\varphi_1 = \varphi', \tau_1 \mid \int_{\tau_2}^{\tau_1} \Delta L d^4x \mid \varphi_2 = \varphi'', \tau_2\rangle \tag{1.7}$$

and

$$\Delta_L\varphi(x) - \Delta^*\varphi(x) = i\left[\varphi(x), \int_{\tau_0}^{\tau} \Delta L \, d^4x + F(\tau)\right], \tag{1.8}$$

where τ is put equal to $n_\mu x^\mu$, and τ_0 corresponds to a space-like hyper-plane σ_0 where all the variations can be considered to vanish. In what follows, τ_0 will be put equal to $-\infty$ from the physical consideration (the so-called adiabatic switch-off assumption). The eq. (1·7) can be directly derived by taking only the variation of Lagrangian, whereas (1·8) by taking into account the definition (1·2) and (I).

The eq. (1·8) can be rewritten in a more general form

$$\varDelta_L f - \varDelta^* f = i\left[\,f(x),\int_{-\infty}^\tau \varDelta L\, d^4x + F(\tau)\right] \tag{II}$$

by virtue of the fact that (II) can be also satisfied by π in place of f.[7] Here f is an arbitrary rational integral functional of $\varphi(x)$ and $\pi(x)$.

§ 2   Further investigation of eqs. (I) and (II).

In the present paper we are only concerned with the variation of the coupling constant g. In this case $\varDelta L$ turns out to be

$$\varDelta L = \varDelta g\left(\frac{\partial L}{\partial g}\right)_{\dot\varphi\varphi}, \tag{2·1}$$

where $(\partial L/\partial g)_{\dot\varphi\varphi}$ means that in taking a derivative of L with respect to g, φ and $\dot\varphi$ should be kept constant.

The eqs. (1·7) and (II) are written down as

$$\frac{d}{dg}\langle\varphi_1=\varphi',\tau_1\,|\,\varphi_2=\varphi'',\tau_2\rangle = i\langle\varphi_1=\varphi',\tau_1\,|\int_{\tau_2}^{\tau_1}\left(\frac{\partial L}{\partial g}\right)_{\dot\varphi\varphi} d^4x\,|\,\varphi_2=\varphi'',\tau_2\rangle \tag{2·2}$$

and

$$\frac{dF}{dg} = \left(\frac{\partial F}{\partial g}\right)_{\varphi\pi} + i\int_{-\infty}^\infty\left[F(\tau),\left(\frac{\partial L(x')}{\partial g}\right)_{\dot\varphi\varphi}\right]\theta(\tau-\tau')d^4x', \tag{2·3}$$

where F is on arbitrary functional of φ and π at a time τ and may explicitly depend on the coupling constant g. (2·3) can be rewritten in the form

$$\frac{dF}{dg} = \left(\frac{\partial F}{\partial g}\right)_{\varphi\pi} - i\int_{-\infty}^\infty\left[F(\tau),\left(\frac{\partial H(x')}{\partial g}\right)_{\varphi\pi}\right]\theta(\tau-\tau')d^4x' \tag{2·3'}$$

by using the well-known relation

$$H = -n^\mu n_\nu T.{}^\nu_\mu = \pi\dot\varphi - L$$

and

$$\left(\frac{\partial H}{\partial g}\right)_{\varphi\pi} = \pi\left(\frac{\partial\dot\varphi}{\partial g}\right)_{\varphi\pi} - \left(\frac{\partial L}{\partial g}\right)_{\dot\varphi\varphi} - \left(\frac{\partial L}{\partial\dot\varphi}\right)\left(\frac{\partial\dot\varphi}{\partial g}\right)_{\varphi\pi} = -\left(\frac{\partial L}{\partial g}\right)_{\dot\varphi\varphi} \tag{2·4}$$

According to the definition of $|\,\varphi(\vec{x},\tau)=\varphi',\tau\rangle$ we have

$$\varphi(\vec{x},\infty)|\,\varphi(\vec{x},\infty)=\varphi',\tau=\infty\rangle = \varphi'(\vec{x})\,|\,\varphi(\vec{x},\infty)=\varphi',\tau=\infty\rangle.$$

By putting

$$S^{-1}\varphi^{in}(\vec{x},\infty)S = \varphi(\vec{x},\infty), \tag{2·5}$$

the above equation is transformed into

$$\varphi^{in}(\vec{x}, \infty)\Psi' = \varphi'(\vec{x})\Psi'$$

$$\Psi' = S \mid \varphi(\vec{x}, \infty) = \varphi', \infty \rangle.$$

Thus we see that Ψ' is an eigenfunction of $\varphi^{in}(\vec{x}, \infty)$ corresponding to a set of eigenvalues $\varphi'(\vec{x})$. Therefore Ψ' can be put equal to

$$\Psi' = S \mid \varphi(\vec{x}, \infty) = \varphi', \infty \rangle = \mid \varphi^{in}(\vec{x}, \infty) = \varphi', \infty \rangle. \qquad (2\cdot6)$$

Here we have assumed that there exists a non-singular matrix S, the unitary character of which will be proved in the later paragraph. By taking into account the equation $(2\cdot6)$, and putting τ_1 and τ_2 in $(2\cdot2)$ equal to $+\infty$ and $-\infty$, respectively, $(2\cdot2)$ turns out to be

$$\frac{d}{dy}\langle \varphi^{in}(\vec{x}, \infty) = \varphi', \infty \mid (S^{-1})^{\dagger} \mid \varphi^{in}(\vec{x}, -\infty) = \varphi'', -\infty \rangle$$

$$= i\langle \varphi^{in}(\vec{x}, \infty) = \varphi', \infty \mid (S^{-1})^{\dagger} \int_{-\infty}^{\infty} \left(\frac{\partial L}{\partial y}\right)_{\varphi\dot{\varphi}} d^4x \mid \varphi^{in}(\vec{x}, -\infty) = \varphi'', -\infty \rangle,$$

$$(2\cdot2)'$$

where $\varphi(\vec{x},\tau)$ has been assumed to tend to $\varphi^{in}(\vec{x}, -\infty)$ in the limit $\tau \to -\infty$ by virtue of the adiabatic procedure. The symbol $\dagger$ means the hermite conjugate of an original operator.) Now we may suppose that $\varphi^{in}(x)$ is independent of y because φ^{in} is a field operator satisfing the equation of free fields, and the initial value of which can be arbitrarily chosen. As a result of this assumption $(2\cdot2)'$ can be rewritten in a form of operator equations as

$$\frac{d(S^{-1})^{\dagger}}{dy} = i(S^{-1})^{\dagger} \int_{-\infty}^{\infty} \left(\frac{\partial L}{\partial y}\right)_{\varphi\dot{\varphi}} d^4x. \qquad (2\cdot2)''$$

Since it is easily seen from $(2\cdot2)''$ that S is a unitary operator provided that L is hermitian, we can further rewrite $(2\cdot2)''$ in the following way

$$\frac{dS}{dy} = iS \int_{-\infty}^{\infty} \left(\frac{\partial L}{\partial y}\right)_{\varphi\dot{\varphi}} d^4x \qquad (2\cdot7)_L$$

or by using the relation $(2\cdot4)$, $(2\cdot7)_L$ runs

$$\frac{dS}{dy} = -iS \int_{-\infty}^{\infty} \left(\frac{\partial H}{\partial y}\right)_{\varphi\pi} d^4x. \qquad (2\cdot7)_H$$

So far all the quantities have been considered as functionals of φ and π (except $(2\cdot7)_L$). In order to write down all the expressions in a covariant form, i. e., in a form independent of n^{μ}, it seems for us more convenient to take φ and $\dot{\varphi}$ as a set of independent field-variables in place of φ and π, respectively. For this purpose let us introduce the symbol θ^* defined as follows:

$$[\varphi(x), \varphi(\tau')]\,\theta^*(\tau-\tau') = [\varphi(\tau), \varphi(\tau')]\,\theta(\tau-\tau'),$$

$$\left[\frac{d}{d\tau}\varphi(\tau), \varphi(x')\right]\theta^*(\tau-\tau') = \frac{d}{d\tau}\left([\varphi(\tau), \varphi(\tau')]\,\theta(\tau-\tau')\right)$$

and $\hspace{10cm}$ (2·8)

$$\left[\frac{d}{d\tau}\varphi(\tau), \frac{d}{d\tau'}\varphi(\tau')\right]\theta^*(\tau-\tau') = \frac{d^2}{d\tau d\tau'}\left([\varphi(\tau), \varphi(\tau')]\,\theta(\tau-\tau')\right).$$

Here it must be remarked that (2·8) is only applicable to the case where the equations of φ are those with the second rank of derivatives with respect to τ. If the rank of derivatives of φ's in the bracket in (2·8) is equal to or higher than that of the equation of fields, we must transform these quantities into those having derivatives of lower rank by means of the equation of fields. Similar consideration can be applied to general cases. By using this symbol and taking into account the Heisenberg equations and (2·4), (2·3) can be written down in the following form :

$$\frac{dF}{dg} = \left(\frac{\partial F}{\partial g}\right)_{\varphi\dot\varphi} + i\int_{-\infty}^{\infty}\left[F(\tau), \left(\frac{\partial L(x')}{\partial g}\right)_{\varphi\dot\varphi}\right]\theta^*(\tau-\tau')\,d^4x', \qquad (2\cdot3)_L$$

where F is an arbitrary functional of φ, $\dot\varphi$ and y. This equation can be easily confirmed by considering two special cases, i. e., $F=\varphi$ and $F=\dot\varphi$. As to the latter case the proof is given as follows :

$$\frac{d\dot\varphi}{dy} = \left(\frac{\partial\dot\varphi}{\partial y}\right)_{\varphi\pi} + i\int\left[\dot\varphi(x), \left(\frac{\partial L(x')}{\partial y}\right)_{\varphi\dot\varphi}\right]\theta(\tau-\tau')\,d^4x'$$

$$= i\left[\left(\frac{\partial\bar{H}}{\partial g}\right)_{\varphi\pi}, \varphi\right] + i\int\frac{d}{d\tau}\left\{\left[\varphi, \left(\frac{\partial L}{\partial g}\right)_{\varphi\dot\varphi}\right]\theta\right\}d^4x'$$

$$- i\int\left[\varphi(x), \left(\frac{\partial L(x')}{\partial y}\right)_{\varphi\dot\varphi}\right]\partial(\tau-\tau')\,d^4x',$$

the first and last terms cancelling with each other by virtue of (2·4), thus we have

$$\frac{d\dot\varphi}{dg} = i\int\left[\dot\varphi(x), \left(\frac{\partial L(x')}{\partial g}\right)_{\varphi\dot\varphi}\right]\theta^*(\tau-\tau')\,d^4x'$$

on account of the definition (2·8).

In the above argument we have assumed the consistency between the two variations, one of them being the translation of space-time coordinates and the other the variation of the coupling constant. The sufficient proof of this consistency is to verify the commutability of $d/d\tau$ and d/dy. $[d/d\tau, d/dy]$ vanishes as can be easily proved.

In concluding this section, it must be noted that all the equations so far obtained can be derived also from the usual Tomonaga-Schwinger's theory. However, we have derived these results without using the usual theory, because we want to apply them to the system with non-local interactions where the usual theory is considered not to be applicable.

§ 3 　Construction of S matrix[8]

In the preceding section the following equations have been derived :

$$\frac{dF}{dg} = \left(\frac{\partial F}{\partial g}\right)_{\varphi\dot\varphi} + i\int_{-\infty}^{\infty}\left[F(\tau), \left(\frac{\partial L(x')}{\partial g}\right)_{\varphi\dot\varphi}\right]\theta^*(\tau-\tau')d^4x', \tag{2.3}_L$$

$$\frac{dS}{dg} = iS\int_{-\infty}^{\infty}\left(\frac{\partial L}{\partial g}\right)_{\varphi\dot\varphi}d^4x, \tag{2.7}_L$$

$$\frac{dF}{dg} = \left(\frac{\partial F}{\partial g}\right)_{\varphi\pi} - i\int_{-\infty}^{\infty}\left[F(\tau), \left(\frac{\partial H(x')}{\partial g}\right)_{\varphi\pi}\right]\theta(\tau-\tau')d^4x' \tag{2.3}_H$$

and

$$\frac{dS}{dg} = -iS\int_{-\infty}^{\infty}\left(\frac{\partial H}{\partial g}\right)_{\varphi\pi}d^4x. \tag{2.7}_H$$

The first two equations are those in Lagrangian formalism, whereas the others in Hamilton formalism.

In this section we shall show how to construct S-matrix by integrating these two sets of equations under the assumption that Lagrangian density has the form

$$L = L_0 + L^I = L_0 + gL^A + g^2L^B$$

Here L_0, L^A and L^B do not explicitly depend on g, and L^I may contain derivatives of field operators, the rank of which is lower than that of L_0. Let us obtain the S-matrix in a form of Taylor series with respect to g. For this purpose the quantity $(d S^N / d g^N)_{g\neq 0}$ is of importance, which results in

$$\left(\frac{d^N S}{dg^N}\right)_{g\neq 0} = \sum_{N/2 \geq m \geq 0}\frac{(i)^{N-m}N!}{m!(N-2m)!} S\int T^*(L_1^B\cdots L_m^B, L'_{m+1}\cdots L'_{N-m})d^4x_1\cdots d^4x_{N-m} \tag{3.1}$$

after a somewhat tedious consideration, which will be explained later. Here L_k and L_k' stand for $L(x_k)$ and $(\partial L(x_k)/\partial g)_{\varphi}$, respectively. T^*-symbol is defined by the following recurrsion formula ;

$$T^*(L_1L_2\cdots L_nL_{n+1}) = L_{n+1}T^*(L_1\cdots L_n) + [T^*(L_1\cdots L_n), L_{n+1}]\theta^* \tag{3.2}$$

where the T^* in the bracket in right-hand side contains several multiple commutators, and in computing these expressions they must be decomposed from inner to outer side. The notation $[\quad]_{\theta^*}$ is defined as follows :

　i) When $A(x)B(x')C(x'')\cdots$ are any set of single operators in Heisenberg representation, it runs

$$[A(x), B(x')]_{\theta^*} = [A(x), B(x')]\theta^*(\tau-\tau').$$

　ii) $[A(x_1)B(x_2), C(x_3)D(x_3)]_{\theta^*}$

$$= [A(x_1), C(x_3)D(x_3)]\theta^*(\tau_1-\tau_3)B(x_2) + A(x_1)$$

$$\cdot[B(x_2), C(x_3)D(x_3)]\theta^*(\tau_2-\tau_3). \tag{3.3}$$

More complicated cases can be similarly computed for. These rules of calculation result from the following fact

$$\frac{d\{F(\tau)G(\tau')\}}{dg} = \left(\frac{\partial(FG)}{\partial g}\right)_{\varphi\dot\varphi} + i\int\left[FG, \left(\frac{\partial L(x'')}{\partial g}\right)_{\varphi\dot\varphi}\right]_{0*} d^4x''. \qquad (3\cdot4)$$

It can be proved that the definition of T^* is equivalent to the ordinary one (c. f. Appendix I).

The proof of $(3\cdot1)$ is given as follows.

Proof : We shall derive the expression $(3\cdot1)$ by means of the method of mathematical induction. In the first place, when $N=2$, we can easily get the relation $(3\cdot1)$ by using $(2\cdot3)$ and $(2\cdot7)$. In the second place let us assume that $(3\cdot1)$ holds for a particular N. Then,

$$\frac{d^{N+1}S}{dg^{N+1}} = \frac{d}{dg}\left(\frac{d^N S}{dg^N}\right) = \sum_{N/2\geq m\geq0}\frac{(i)^{N-m}N!}{m!(N-2m)!}$$

$$\cdot\left\{\frac{dS}{dg}\int T^*(L_1^B\cdots L_m^B L'_{m+1}\cdots L'_{N-m})d^4x_1\cdots d^4x_{N-m}\right.$$

$$\left. + S\int\frac{d}{dg}T^*(L_1^B\cdots L_m^B L'_{m+1}\cdots L'_{N-m})d^4x_1\cdots d^4x_{N-m}\right\}.$$

Substituting with $(2\cdot7)_L$ into the expression of the righthand side, and taking account of $(3\cdot4)$, we have

$$= \sum_{N/2\geq m\geq0}\frac{(i)^{N-m}N!}{m!(N-2m)!}\left\{iS\int L'_{N+1-m}T^*(L_1^B\cdots L_m^B L'_{m+1}\cdots L'_{N-m})d^4x_1\cdots d^4x_{N-m+1}\right.$$

$$+ iS\int[T^*(L_1^B\cdots L_m^B L'_{m+1}\cdots L'_{N-m}), L'_{N+1-m}]_{0*} d^4x_1\cdots d^4x_{N+1-m}$$

$$+ S2(N-2m)\int T^*(L_1^B\cdots L_m^B L_{m+1}^B L'_{m+2}\cdots L'_{N-m})d^4x_1\cdots d^4x_{N-m}.$$

By using the definition of $(3\cdot2)$, it runs

$$= \sum_{N/2\geq m\geq0}\frac{(i)^{N+1-m}N!}{m!(N-2m)!}S\int T^*(L_1^B\cdots L_m^B L'_{m+1}\cdots L'_{N+1-m})d^4x_1\cdots d^4x_{N+1-m}$$

$$+ \sum_{N/2\geq m\geq0}\frac{(i)^{N-m}N!2(N-2m)}{m!(N+1-2m)!}S\int T^*(L_1^B\cdots L_{m+1}^B L'_{m+2}\cdots L'_{N-m})d^4x_1\cdots d^4x_{N-m}$$

or, by changing the labels of L's, it results in

$$= \sum_{N+1/2\geq m\geq0}\frac{(i)^{N+1-m}N!(N+1-2m)}{m!(N+1-2m)!}S\int T^*(L_1^B\cdots L_m^B L'_{m+1}\cdots L'_{N+1-m})d^4x_1\cdots d^4x_{N+1-m}$$

$$+ \sum_{N+1/2\geq m\geq0}\frac{(i)^{N+1-m}N!2(N+2-2m)}{(m-1)!(N+2-2m)!}S\int T^*(L_1^B\cdots L_m^B L'_{m+1}\cdots L'_{N+1-m})d^4x_1\cdots d^4x_{N+1-m}$$

$$= \sum_{N+1/2\geq m\geq0}\frac{(i)^{N+1-m}(N+1)!}{m!(N+1-2m)!}S\int T^*(L_1^B\cdots L_m^B L'_{m+1}\cdots L'_{N+1-m})d^4x_1\cdots d^4x_{N+1-m}.$$

Q.E.D.

Putting q equal to zero in $(3\cdot1)$, we get the expression of S in a form of power series in q as follows,

$$S=\sum_{N=0}^{\infty}\sum_{N/2\geq m\geq0}\frac{q^N(i)^{N-m}}{m!(N-2m)!}\iint T^*(L_1^{R\ in}\cdots L_m^{R\ in}L_{m+1}^{A\ in}\cdots L_{N-m}^{A\ in})d^4x_1\cdots d^4x_{N-m}$$

$$=\sum_{n=0}^{\infty}\frac{(i)^n}{n!}\int T^*(L_1^{I\ in}L_2^{I\ in}\cdots L_n^{I\ in})d^4x_1\cdots d^4x_n, \tag{$3\cdot5$}_L$$

where L^{in} means the interaction operator.

In Hamilton formalism, the expression of S-matrix can be obtained in quite a similar way, the result of which is

$$S=\sum_{n=0}^{\infty}\frac{(-i)^n}{n!}\int T(H_1^{I\ in}\cdots H_n^{I\ in})d^4x_1\cdots d^4x_n. \tag{$3\cdot5$}_H$$

Thus it has been seen that the difference between these two expressions $(3\cdot5)_L$ and $(3\cdot5)_H$ arises from which set of field quantities we take as independent variables.

It seems to be worth-while to show another method of construction of S-matrix which is given by direct integration of $(2\cdot7)_L$, provided that the q-dependence of Heisenberg operators may be known. The expression of S thus obtained is as follows,

$$S=\sum_{n=0}^{\infty}(i)^n\int_0^q dq_1\int_0^q dq_2\cdots\int_0^q dq_n\ \bar{L}'_n\bar{L}'_{n-1}\cdots\bar{L}'_1\theta(q-q_1)\theta(q_1-q_2)\cdots\theta(q_{n-1}-q_n),$$

where the abbreviation $\bar{L}'_k=\int d^4x_k\left(\dfrac{\partial L_k}{\partial q}\right)_{\varphi\dot{\varphi}}$ has been used. This expression seems very convenienient in applying it to non-local cases, because we can hardly expect that $(2\cdot3)_L$ can be made to hold in such a case.

§4   Application to the case of non-local interaction[9]

When one applies Schwinger's formalism to the non-local case, it is necessary to exclude some parts of fundamental postulates which are allowable only in the local case. By virtue of a good nature of the form factor mentioned in the introduction, state vectors may be defined only in the remote past and future as a functional of $\varphi^{in}(\vec{x},\tau=-\infty)$ and $\varphi^{out}(\vec{x},\tau=+\infty)$ respectively. Therefore it would be reasonable to adopt the following fundamental postulates :

i)$_N$ State vectors

$$|\varphi^{in}(\vec{x},\tau=-\infty)=\varphi'(\vec{x}),\tau=-\infty\rangle,\ |\varphi^{out}(\vec{x},\tau=+\infty)=\varphi'(\vec{x}),\tau=+\infty\rangle,$$

and transformation function

$$\langle\varphi^{out}(\vec{x},\tau=+\infty)=\varphi'(\vec{x}),\ \tau=+\infty\ |\ \varphi^{in}(\vec{x},\tau=-\infty)=\varphi''(\vec{x}),\ \tau=-\infty\rangle$$

exist. Here $\varphi^{out}(x)$ is a field operator satisfying the equation of free-fields and is supposed to agree with $\varphi(x)$ at the limit $\tau\rightarrow+\infty$ under the assumption of the adiabatic procedure.

ii)$_N$ Matrix elements

$$\langle \varphi^{in}(\vec{x}, \tau=-\infty)=\varphi', \tau=-\infty \mid A(x) \mid \varphi^{in}(\vec{x}, \tau=-\infty)=\varphi'', \tau=-\infty \rangle$$

$$\text{(4·1)}$$

$$\langle \varphi^{out}(\vec{x}, \tau=+\infty)=\varphi', \tau=\infty \mid A(x) \mid \varphi^{in}_{out}(\vec{x}, \tau=\mp\infty)=\varphi'', \tau=\mp\infty \rangle$$

can be defined. In the local case these quantities can be computed by means of the postulate iv) and the initial condition ii). Accordingly in order to compute these quantities a new condition is required which serves for the initial condition ii). As to the computation of these quantities we shall show it later.

iii)$_N$ All quantities like (4·1) (in case of $A=1$, it is just the transformation function) satisfy the combination law of matrices. Therefore the second quantities of (4·1) can be derived from i)$_N$ and the first quantity of (4·1).

iv)$_N$ Let us postulate Schwinger's equation

$$\varDelta\langle \varphi^{out}(\vec{x}, \tau=+\infty)=\varphi', \tau=+\infty \mid \varphi^{in}(\vec{x}, \tau=-\infty)=\varphi'', \tau=-\infty \rangle$$

$$\text{(4·2)}$$

$$=i\langle \varphi^{out}(\vec{x}, \tau=\infty)=\varphi', \tau=+\infty \mid \varDelta I \mid \varphi^{in}(\vec{x}, \tau=-\infty)=\varphi'', \tau=-\infty \rangle,$$

where the variations should be restricted so as to have no relations with local concepts. In the local case, we computed I by assigning arbitrary C numbers $\varphi''(x)$ at every point x in such a way as being continuously differentiable all over the domain. Since this procedure has no relations with field equations, we can take a small arbitrary change of assignment

$$\varphi^{H}(x)\rightarrow\varphi^{H}(x)+\delta^{*}\varphi(x) \quad \delta^{*}\varphi^{H}(x)=0 \ at \pm\infty,$$

by virtue of which Euler equations can be derived as in the local case. In the second place, we can take the variation $g\rightarrow g+\varDelta g$, under which the interaction Lagrangian changes uniformly all over the world, since this concept is independent on local character.

From this variation we have

$$\frac{d}{dg}\langle \varphi^{out}(\vec{x}, \tau=+\infty)=\varphi', \tau=+\infty \mid \varphi^{in}(\vec{x}, \tau=-\infty)=\varphi'', \tau=-\infty \rangle$$

$$=i\langle \varphi^{out}=\varphi', \tau=+\infty \mid \int_{-\infty}^{\infty}\left(\frac{\partial L(\xi)}{\partial g}\right)d\xi \mid \varphi^{in}=\varphi'', \tau=-\infty \rangle, \quad \text{(4·3)}$$

where ξ represents a set of coordinates variables which appear in the interaction Lagrangian. φ^{in} has been assumed to be independent on g, while φ^{out} depends on g, because the latter is a functional of φ^{in} and g, whose explicit expression was given in Bloch's paper.[3] $\partial L/\partial g$ means a derivative with respect to g appearing explicitly in L.

v)$_N$ S-matrix exists, and is defined by the following equation

$$S^{-1}\varphi^{in}(x)S=\varphi^{out}(x) \quad \text{(4·4)}$$

for an arbitrary point x.

We can not determine a transformation function only from the equation (4·3) and the same is true for matrix elements. Therefore we take the following conditions

$$\lim_{\eta \to 0} \langle \varphi^{out} = \varphi', \tau = \infty \mid \varphi^{in} = \varphi'', \tau = -\infty \rangle = \langle \varphi^{in} = \varphi', \infty \mid \varphi^{in} = \varphi'', -\infty \rangle \quad (4\cdot5)$$

and

$$\lim_{\eta \to 0} \varphi(x) = \varphi^{in}(x). \quad (4\cdot6)$$

Bloch's expression of φ is just the solution of Euler's equations under the condition $(4\cdot6)$. As a result of these considerations matrix elements

$$\langle \varphi^{in} = \varphi', \tau = -\infty \mid \varphi(x) \mid \varphi^{in} = \varphi'', \tau = -\infty \rangle$$

can be evaluated and the postulate ii)$_N$ turns out to be well-defined.

As in § 2, the transformation function can be written in the following way;

$$\langle \varphi^{out} = \varphi', \tau = +\infty \mid \varphi^{in} = \varphi'', \tau = -\infty \rangle$$
$$= \langle \varphi^{in} = \varphi', +\infty \mid (S^{-1})^{\dagger} \mid \varphi^{in} = \varphi'', \tau = -\infty \rangle, \quad (4\cdot7)$$

owing to the postulate (v)$_N$. Accordingly $(4\cdot3)$ can be rewritten as

$$\frac{d(S^{-1})^{\dagger}}{dg} = i(S^{-1})^{\dagger} \int \frac{\partial L}{\partial g} d\xi,$$

or

$$\frac{dS}{dg} = iSW, \quad (4\cdot8)$$

where $W = \int (\partial L / \partial g) d\xi$. The last equation arises from the fact that S is easily seen to be unitary if the form factors satisfy the hermitian condition.

From $(4\cdot8)$ and $(4\cdot4)$ we get the following equation,

$$\frac{d\varphi^{out}(x)}{dg} = i \left[\varphi^{out}(x), W \right]. \quad (4\cdot9)$$

If both sides of this equation are computed separately and turn out not to coincide with each other, at least one of two assumptions $(4\cdot4)$ or $(4\cdot8)$ must be incorrect. In order to investigate this situation, we shall introduce the following notations.

a) Let $A, B, \cdots$ be single operators in Heisenberg representation. $[A, B]'$ is meant by

$$'[A, B] = [A, B]' \equiv \lim_{g \to 0} [A, B]$$
$$'\{A, B\} = \{A, B\}' \equiv \lim_{\eta \to 0} \{A, B\}$$

a)' $\quad '[A, B]_{\theta^*} = [A, B]'_{\theta^*} = \lim_{g \to 0} [A, B] \theta(\tau_A - \tau_B)$

b) $\quad '[AB, CD] = '[A, CD] B + A '[B, CD]$
$$= '[A, C] DB + C'[A, D] B + A'[B, C] D + AC'[B, D]$$
$$[AB, CD]' = [AB, C]' D + C [AB, D]'$$
$$= [A, C]' BD + A [B, C]' D + C [A, D]' B + CA [B, D]'$$

b)′ $'[AB, CD]_0* = DB\,'[A,C]_0* + CB'[A,D]_0* + AD'[B,C]_0* + AC'[B,D]_0*$,

$[AB, CD]'_0* = BD[AC]'_0* + AD[B,C]'_0* + CB[A,D]'_0* + CA[B,D]'_0*$.

It is easily seen how to decompose more general expression such as

$[AB\cdots, MN\cdots]'$, $'[AC\cdots, MN\cdots]$, $[AB\cdots, MN\cdots]'_0*$ and $'[AB\cdots, MN\cdots]_0*$.

By using these notations, the integral equations satisfied by φ and φ^{out} are written as

$$\varphi(x) = \varphi^{in}(x) + ig\,'[\varphi(x), W]_0* , \qquad (4\cdot10)_a$$

$$\varphi(x) = \varphi^{out}(x) - ig'[\varphi(x), W]_0* , \quad \theta^*(\tau) = \theta^*(-\tau), \qquad (4\cdot10)_b$$

$$\varphi^{out}(x) = \varphi^{in}(x) + ig\,'[\varphi(x), W] . \qquad (4\cdot10)_c$$

Differentiating $(4\cdot10)_a$ with respect to g we obtain the following equation

$$\frac{d\varphi(x)}{dg} = i'[\varphi(x), W]_0* + ig\frac{d}{dg}\,'[\varphi(x), W]_0*$$

where d/dg in the right hand side means that we must differentiate only those operators which originally belonged to W and which were remained after computing the commutator. We shall denote this operator with $\overset{\circ}{d}/dg$.

Similarly we get

$$\frac{df}{dg} = i'[f, W]_0* + ig\frac{\overset{\circ}{d}}{dg}'[f, W]_0* \qquad (4\cdot11)$$

where f is a polynomial of φ and $\varphi_{,\mu}$, and explicitly independent on g. Repeating this operation we obtain

$$\frac{\overset{\circ}{d}}{dg_1}'[\varphi, W_1]_0* = i'['[\varphi, \overline{W_1}]_0* W_2]_0* + ig\frac{\overset{\circ}{d}}{dg_2}'['[\varphi, \overline{W_1}]_0* W_2]_0* .$$

According to the definition of $\overset{\circ}{d}/dg_1$ in the left hand side of this equation, operators of W_2 are considered as non-commutative only with operators in W_1. This fact is shown by symbol ⌐ in the above equation. $\overset{\circ}{d}/dg_n$ means that operators in W_n must be only differentiated after computation of the commutator. In consequence of this operation the following equation can be easily derived

$$\frac{df}{dg} = i\sum_{n=0}^{\infty} (ig)^{n'}[\cdots'[\overline{f, W}]_0* \overline{W_1}]_0* \cdots \overline{W_n}]_0* \qquad (4\cdot12)$$

where multiple commutator should be computed from inner to outer and W_k is considered to commute all others except W_{k-1}. Starting from $(4\cdot10)_c$ we get in a similar manner

$$\frac{d\varphi^{out}(x)}{dg} = i\sum_{n=0}^{\infty} (ig)^{n'}[\cdots'[\overline{\varphi(x)\,W}]\,\overline{W_1}]_0* \cdots \overline{W}]_0* \equiv iA(x). \qquad (4\cdot13)$$

For computing $[\varphi^{out}(x), W]$ we consider at first

$$[\varphi^{out}(x), \varphi_1] = [\varphi^{out}, \varphi_1] - ig\, [\varphi^{out} \,'[\varphi_1, W]_0^*].$$

By using the relation $[\varphi^{out}, \varphi^{out}] = [\varphi^{in}, \varphi^{in}] = \lim_{y \to 0} [\varphi, \varphi]$ (on account of v)$_N$) it reads

$$[\varphi^{out}(x), \varphi_1] = [\varphi, \varphi_1]' - ig\, [\varphi^{out} \,'[\varphi_1, W]_0^*],$$

or more generally we have

$$[\varphi^{out}(x), f] = [\varphi, f]' + ig\, [\overline{\varphi^{out} [W, f]}\,'_0^*]. \qquad (4\cdot14)$$

Since the second term of the right hand side can be written as

$$[\overline{\varphi^{out}[W, f]}\,'_0^*] = [\,\overline{[\varphi^{out}, W]f}\,'_0^*$$

$(4\cdot14)$ becomes

$$[\varphi^{out}, f] = [\varphi, f]' + ig\, [\,[\overline{\varphi_{out}, W]f}\,'_0^*. \qquad (4\cdot15)$$

Repeating this operation we obtain the following expression

$$[\varphi^{out}, f] = [\varphi, f]' + \sum_{n=0}^{\infty}(ig)^n[\cdots[\overline{\varphi^{out}, W_1}\,'_0^* \, \overline{W_2}\,'_0^* \cdots \overline{W_n}\,'_0^* f\,]'_0^*.$$

Taking $f = W$, it becomes

$$[\varphi^{out}, W] = \sum_{n=0}^{\infty}(ig)^n[\cdots[\overline{\varphi^{out}, W}\,]'\, \overline{W_1}\,'_0^* \cdots \overline{W_n}\,]'_0^* \qquad (4\cdot16)$$

where multiple commutator should be computed in a way similar to $(4\cdot13)$ except the different order of computation concerning the primed symbol. $(4\cdot9)$ requires that A is equal to B, but unfortunately we can not expect this equality because $[\ ]'$ and $'[\ ]$ are different in case of a general form factor. In the appendix II we shall show this situation in the system whose interaction Lagrangian density is

$$L' = gF(xx'x'')\overline{\varphi}(x)u(x')\varphi(x''). \qquad (4\cdot17)$$

There we shall see $(4\cdot10)$ breaks down in the third order of g. Thus we have reached the conclusion that at least one of two assumptions, $(4\cdot4)$ and $(4\cdot8)$, is wrong if no special condition is imposed on the form factor.

Bloch told before the International Congress in Kyoto that, in the above example, S-matrix satisfying $(4\cdot4)$ can not be found in the 4th order of g.[6] This fact seems to be equivalent to the breaking down of $(4\cdot9)$ in the 3rd order of g in our case. Even in the local case, we can hardly show the equality between A and B, but in this case it can be proved from other side (§ 2). It is shown in the appendix II that the difference between A and B in the 3rd order or g has closed chain of θ-functions and hence vanishes in the local case. Thus the condition on form factors which ensures $(4\cdot9)$ would be very severe. We know little as to this condition. It seems that Bloch did not use this condition explicitly in his formal proof of the existence of S-matrix. In the present stage it is most desirable to investigate what a kind of mathematical condition on

form factors will ensure the existence of S-matrix, since it is not obvious whether the condition of " normal class " is equivalent or not to that mentioned above, and further the mathematical definition itself of " normal class " has not yet been clarified.

If we can find fortunately a particular form factor which permits the existence of S-matrix, a particular hermitian operator H must exist by means of which the equation of S-matrix ean be written down as

$$\frac{dS}{dg} = iSH.$$

Though it is not sure that this H must be equal to W, we can expect the identity between H and W in the non-local case since W is an invariant integral all over the space time.

Acknowledgments

We would like to express our gratitude to Professor H. Yukawa for his kind hospitality and valuable discussions during our stay at Yukawa Hall.

Appendix I

Let us denote tentatively the chronological operator introduced in $(3 \cdot 4)$ with $T^{*\prime}$, whereas the ordinary one with T^{*}. Namely, $T^{*\prime}$ is defined by

$$T^{*\prime}_{n+1} = L_{n+1} T^{*\prime}_n + [T^{*\prime}_n, L_{n+1}]_0^*$$

and

$$T_n^{*\prime} = T^{*\prime}(L_1 \cdots L_n),$$

while T^* by

$$T_n^* = T^*(L_1 \cdots L_n) = \sum_{perm(n)} L_1 \cdots L_n \theta^*(1 \cdots n),$$

where the summation must be carried out over all the possible permutation of $(1, 2, \cdots n)$ and $\theta^*(1, 2, \cdots n)$ is meant by

$$\theta^*(1, \cdots n) = \theta^*(1, 2)\theta^*(2, 3) \cdots \theta^*(n-1, n)$$

and

$$\theta^*(i, k) = \theta^*(\tau_i - \tau_k).$$

We shall give a proof of the equivalence between $T^{*\prime}$ and T^* by virtue of the method of mathematical induction. In the first place when $n=1$ it is easily seen that

$$T_2^{*\prime} = L_2 L_1 + [L_1, L_2]_0^* = T_2^*.$$

In the second place, let us assume that $T_n^{*\prime}$ is identical with T_n^* for particular n. Then from this assumption and the definition of $T^{*\prime}$, we get

$$T^{*\prime}_{n+1} = L_{n+1} T_n^* + [T_n^*, L_{n+1}]_n^*.$$

By using the definition of T^*, the right-hand side is transformed into

$$= \sum_{perm(n)} \{L_{n+1}L_1L_2\cdots L_n\theta^*(1\cdots n) + [L_1L_2\cdots L_n, L_{n+1}]_0^*\theta^*(1\cdots n)\},$$

or it can be rewritten as follows,

$$= \sum_{perm(n)} \{L_{n+1}L_1L_2\cdots L_n\theta^*(1\cdots n)(\theta^*(n+1,1) + \theta^*(1,n+1))$$

$$+ \sum_{i=1}^{n} L_1\cdots L_{i-1}[L_i, L_{n+1}]L_{i+1}\cdots L_n\theta^*(i,n+1)\theta^*(1\cdots n)\}.$$

Decomposing the commutators in this expression, and changing the labels of L, we get

$$= \sum_{perm(n)} \{L_{n+1}L_1L_2\cdots L_n\,\theta^*(n+1,1,2\cdots n)$$

$$+ \sum_{i=1}^{n} L_1\cdots L_iL_{n+1}L_{i+1}\cdots L_n(\theta^*(1\cdots i, n+1, i+1, \cdots n) + \theta^*(1\cdots n)\theta^*(i+1,n+1))$$

$$- \sum_{i=1}^{n} L_1\cdots L_iL_{n+1}L_{i+1}\cdots L_n\theta^*(1\cdots n)\theta^*(i+1,n+1)\},$$

where the relation $\theta^*(1\cdots n)\theta^*(i,n+1) = \theta^*(1\cdots i, n+1, i+1, \cdots n) + \theta^*(1\cdots n)\theta^*(i+1, n+1)$ has been used. This expression can be easily transformed into

$$= \sum_{perm(n)} \{L_{n+1}L_1L_2\cdots L_n\theta^*(n+1,1,\cdots n)$$

$$+ \sum_{i=1}^{n} L_1\cdots L_iL_{n+1}L_{i+1}\cdots L_n\theta^*(1\cdots, n+1, i+1, \cdots n)\}$$

$$= \sum_{perm(n+1)} L_1L_2\cdots L_nL_{n+1}\,\theta^*(1\cdots n, n+1)$$

$$= T^*_{n+1}.$$

Appendix II

For analysis of the quantities A and B in case of the interaction Lagrangian density $(4\cdot17)$, the following graphical consideration is very convenient. In our diagram several quantities will be shown as follows:
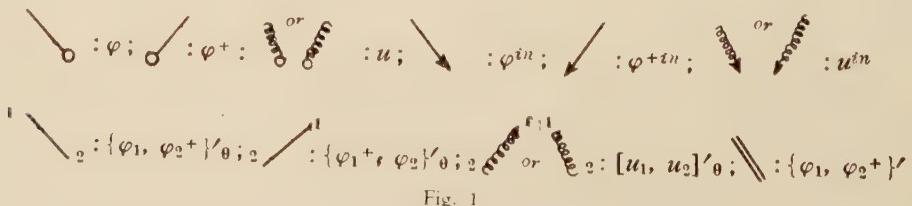

Fig. 1

By using these rules, the term with $n=3$ in A can be represented as
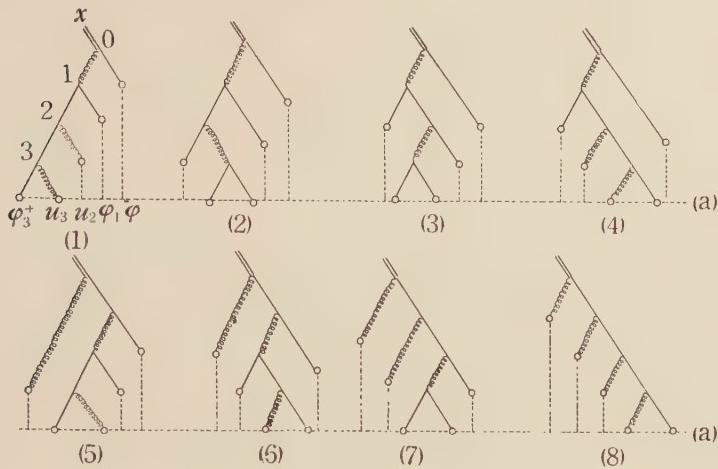
Fig. 2 φ should be read as φ_0

where the order of the projective points on the axis (a) in a ·particular diagram corresponds to that of the Heisenberg operators appearing in a particular term which has been derived from A after the computation of commutators. For instance, the term corresponding to the graph of left end is given by

$$\{\varphi_x\varphi_0{}^+\}'[u_0u_1]'_\theta\{\varphi_1{}^+\varphi_2\}'_\theta\{\varphi_2{}^+\varphi_3\}'_\theta\varphi_3{}^+u_3u_2\varphi_1\varphi_0$$

apart from form factors and a numerical constant. The term with $n=3$ in B can be represented in a similar manner
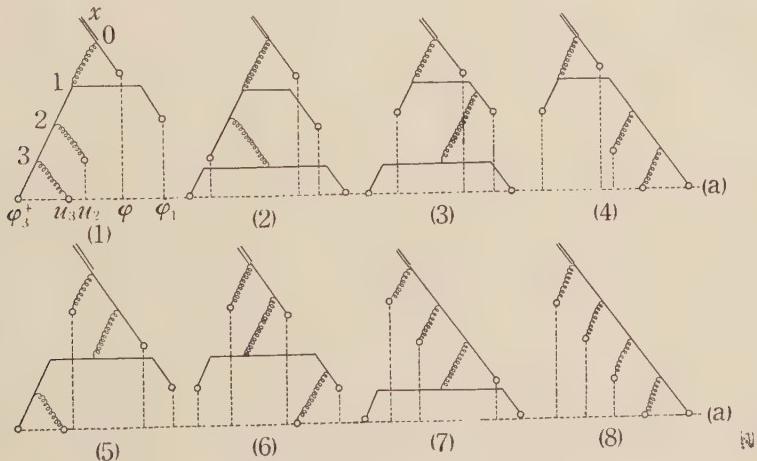
Fig. 3 φ should be read as φ_0

Each Heisenberg operator can be developed in a power series of g, and each term of this series can be represented by a similar graph. For example, the term with g^3 in the expansion of φ can be shown by

$$\varphi_3^{+in} \quad u_1^{in} \quad \varphi_2^{in} \quad \varphi_1^{in}$$
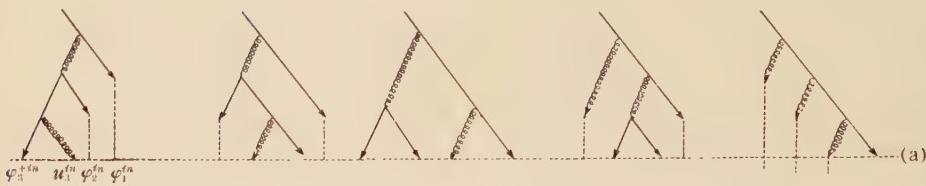
Fig. 4

Therefore we can easily get the expression of any term with arbitrary order in g appearing in A and B, by following these two steps of consideration. By using these methods A and B can be compared graphically, and the difference between them is caused only by the order of $\varphi^{in}\varphi^{+in}u^{in}$. The explicit expression of this difference is in case of the 3rd order in g.

$$\{\varphi_r\varphi_0^+\}_0'\{\varphi_0\varphi_3^+\}'[u_3u_2]'_0\{\varphi_2\varphi_1\}'_0[u_1u_0]'_0\varphi_1^{+in}\varphi_3^{in}\varphi_2^{in}$$

$$- \{\varphi_r\varphi_0^+\}'[u_0u_1]'_0\{\varphi_1\varphi_2^+\}'_0[u_2u_3]'_0\{\varphi_3^+\varphi_0\}'_0\varphi_1^{+in}\varphi_3^{in}\varphi_2^{in}.$$

This vanishes in the local case because it contains the closed chain of θ-function.

References

1) H. Yukawa, Phys. Rev. 77 (1950), 219; 80 (1950), 1047.
2) M. Fierz, ibid. 78 (1950), 184.
3) P. Kristensen and C. Møller, Danske Viedensk. Selsk. Mat.-fys.
    Medd. 27 (1952), No. 7.
    C. Bloch, ibid. 27 (1952), No. 8
    M. Chrétien and R. E Peierls, Nuovo Cimento 10 (1953), 669.
    Y. Takahashi and H. Umezawa, Prog. Theor. Phys. 9 (1953), 14.
    Y. Katayama, ibid. 10 (1953), 31.
4) W. Pauli, Nuovo Cimento 10 (1953), 648.
5) J. Schwinger, Phys. Rev. 82 (1951), 914.
6) In the International Congress in Kyoto Bloch proposed to determine S-matrix in every order of g
    by adding some terms to the original Lagrangian density in such a way as to be able to get the
    S-matrix in that order of g. However, it seems that this procedure results in substituting another
    problem for the given one, and further it is quite doubtful if this procedure will converge.
7) In order to prove the relation (II) in case of $F=\pi$, it is necessary to know the expression of $\Delta\pi$
    as a functional of φ, π and $\Delta\varphi$. $\Delta\pi$ is determined as follows :
    Let it be supposed that π is transformed into $\bar{\pi}=\pi+\Delta\pi$ according to the variation of $\varphi\rightarrow\bar{\varphi}=\varphi+\Delta\varphi$.
    Since π must be canonically conjugate to $\bar{\varphi}$, we have the relation ;

$$[\pi(\dot{x},\tau),\,\varphi(\dot{x}'',\tau)]-[\bar{\pi}(\dot{x},\tau),\,\varphi(\dot{x}'',\tau)]=1\,i\cdot\delta(\dot{x}-\dot{x}'').$$

    If $\Delta\varphi$ is assumed to be commutative with φ, we can get an equation

$$[\Delta\pi,\varphi]+[\pi,\Delta\varphi]=\left(\frac{\delta\Delta\pi}{\delta\pi}+\frac{\delta\Delta\varphi}{\delta\varphi}\right)[\pi,\varphi]=0$$

    from the above relation, By integrating this with respect to, π, $\Delta\pi$ is given as follows,

$$\Delta\pi=-\frac{1}{2}\left\{\pi\frac{\delta\Delta\varphi}{\delta\varphi}+\frac{\delta\Delta\varphi}{\delta\varphi}\pi\right\}.$$

8) W. K. Burton, Phys. Rev. 84 (1951), 159 (L.).
9) A similar work was done by C. Levinson, the preprint of which was shown us by Professor
    H. Yukawa.

Theory of Antiferromagnetic Resonance in CuCl$_2$·2H$_2$O

Takeo NAGAMIYA

Department of Physics, Osaka University

After giving a short review of Nagamiya-Yosida's theory of antiferromagnetic resonance, it is shown that this theory is equivalent to Gorter-Ubbink's at absolute zero, so far as the external field H employed in the former is replaced by H' defined by $2H_x' = g_a H_x$, etc. and the parallel and perpendicular susceptibilities are related with observed susceptibilities in a suitable way (equations (1.9)), where g_a is the value of the g-tensor in the direction of the crystalline a-axis, etc. (§ 1) It is then shown that the small anisotropy of the Weiss molecular field can be taken into account by introducing an anisotropy energy of orthorhombic symmetry and that in this way the anisotropic Weiss field can be replaced by an isotropic one (§ 2). By using this anisotropy energy, the hyperbola in the ac-plane that defines the critical field strength is derived in a simple way for an arbitrary temperature, without making use of Weiss approximation (§ 2). In § 4, then, a summarizing review of Yosida's theory of resonance is given, supplementing it with a few more resonance formulas. In § 5, it is shown that in a certain range of temperature resonance takes place at the critical field strength. There is a case such that the resonance absorption curve is semi-infinitely broad, being cut off for the field strengths inferior to the critical field strength, and this occurs at one end of that temperature range. This is in accord with Ubbink's experiment. In § 6 it is shown that all the observed resonance peak positions obtained by varying the temperature and the direction of the applied field are quantitatively well explained, if one determines the three parameters entering the theory (two anisotropy constants, the ratio between them being independent of temperature, and the ratio between the parallel and perpendicular susceptibilities) from a part of the resonance data and the measurements of the susceptibilities. Some discussions are given about the change of the width of resonance curve with the change of the direction of the applied field and also about the polarization effect. In § 7 are given further predictions from the theory.

§ 1. Introduction

In a series of papers the writer and Yosida have developed a theory of antiferromagnetic resonance phenomenon based on Weiss molecular field model.[1-4] In the first paper of the writer, he assumed a system of classical spins with an anisotropy energy of the form

$$(1/2) \, K [(\beta_+^2 + \gamma_+^2) + (\beta_-^2 + \gamma_-^2)], \qquad (1 \cdot 1)$$

α_+, β_+, γ_+, and α_-, β_-, γ_- being the direction cosines of the + and − spins and K the anisotropy constant, and the equation of motion

$$\boldsymbol{M}^{\pm} = \gamma \, \boldsymbol{M}^{\pm} \times (\boldsymbol{H} - A\boldsymbol{M}^{\pm} + \boldsymbol{H}_a^{\pm}), \quad \gamma = g\mu_b / \hbar \qquad (1 \cdot 2)$$

was solved in the neighbourhood of the equilibrium position to obtain the characteristic frequency. In $(1 \cdot 2)$, $\boldsymbol{M}^{\pm}$ are the magnetic moments of the two sublattices, A the Weiss field constant, $\boldsymbol{H}$ the applied field, and $\boldsymbol{H}_a^{\pm}$ the anisotropy field acting on $\boldsymbol{M}^{\pm}$, which is derived from $(1 \cdot 1)$ as

$$H_a^\pm = -\frac{K}{M_0^2}\,(O,\ M_y^\pm,\ M_z^\pm), \qquad\qquad (1\cdot3)$$

where M_0 is the magnitude of $M^\pm$ for $H=0$.

This theory was further developed by Yosida for spins of finite magnitude and in particular for spins $1/2$. In the latter case, he started from the general expression given by van Vleck[5] for the interaction between two spins S_l and S_m of magnitude $1/2$:

$$-2J_{lm}\sum_{xyz}S_{lx}\,S_{mx}+C_{lm}\,(2S_{lz}\,S_{mz}-S_{lx}\,S_{mx}-S_{ly}\,S_{my})$$
$$+D_{lm}(S_{lx}\,S_{mx}-S_{ly}\,S_{my})$$
$$+E_{lm}(S_{lx}\,S_{my}+S_{ly}\,S_{mx})+F_{lm}(S_{lx}\,S_{mz}+S_{lz}\,S_{mx})$$
$$+G_{lm}(S_{ly}\,S_{mz}+S_{lz}\,S_{my}),$$

where J_{lm}, C_{lm}, D_{lm}, E_{lm}, F_{lm}, and G_{lm} are the coupling constants which depend on the relative positions of the two spins in the crystal. For temperatures below the Néel point he took the Weiss approximation by taking the sum of the above expression over the lattice sites for m which surround the site for l and taking the thermal average over the orientation of S_m. Then, if one confines himself to the case of an orthorhombic crystal (which is the case for $CuCl_2 \cdot 2H_2O$), one obtains the Weiss type Hamiltonian for S_l in the form

$$a_x S_x^+\langle S_x^-\rangle+a_y S_y^+\langle S_y^-\rangle+a_z S_z^+\langle S_z^-\rangle$$
$$+c_x S_x^+\langle S_x^+\rangle+c_y S_y^+\langle S_y^+\rangle+c_z S_z^+\langle S_z^+\rangle. \qquad (1\cdot5)$$

The symbol $\langle\ \rangle$ means the thermal average, those S^+'s which are outside this symbol stand for S_l, which we assume to lie on the $+$ sublattice, and those S^-'s and S^+'s which are inside stand for those S_m which lie on the $-$ and $+$ sublattices, respectively. The x, y, z axes are taken along the crystalline axes, so that non-diagonal terms do not appear.

It can be shown, as already done by Yosida, that, when the anisotropy of the coefficients a_x, a_y, a_z and c_x, c_y, c_z is small, the anisotropy of $(1\cdot5)$ can be removed by introducing the anisotropy energy of the form

$$(1/2)\ K_1(\beta_+^2+\beta_-^2)+(1/2)K_2(\gamma_+^2+\gamma_-^2). \qquad (1\cdot6)$$

(We shall give a simple proof of this in the next section). The corresponding anisotropy field is

$$H_a^\pm = -\left(0,\ \frac{K_1}{M_0^2}\,M_y^\pm,\ \frac{K_2}{M_0^2}\,M_z^\pm\ \right). \qquad (1\cdot7)$$

Therefore, the resonance phenomenon can be treated in a similar way to that mentioned above. Yosida applied his theory to the explanation of the experiments on $CuCl_2 \cdot 2H_2O$ done at Leiden.[6]

In the meantime, Gorter and Haantjes[7] and Ubbink[8] worked out a theory of resonance absorption in orthorhombic crystal which has much in common with Yosida's

theory. Although they confined themselves to absolute zero, they made a very detailed analysis of the magnitudes and directions of the spin vectors of the two sublattices and the frequencies of the characteristic oscillations as functions of applied magnetic field. They stressed in particular the importance of the anisotropic character of the spin magnetic moment, i. e., the tensor character of the *g*-factor, as well as the anisotropic character of the Weiss field. Yosida did not overlooked these facts in computing the susceptibility above the Néel point (indeed, he did not use Weiss approximation for this case), but simply neglected the tensor character of g below the Néel point. Gorter and his collaborators have furthermore found an interesting fact that there exists a hyperbola (which we shall call *critical hyperbola*) in the plane containing the easiest axis of magnetization (the a-axis in the case of CuCl₂·2H₂O) and the most difficult axis (the c-axis in the same case) such that, if the strength of the magnetic field applied in this ac-plane exceeds a certain critical value defined by this hyperbola, the common axis of the antiparallel spins turns from the neighbourhood of the a-axis to the b-axis. This is the generalization of Néel's original theory[9] in the case of uniaxial anisotropy for absolute zero which had been generalized by Yosida[2] for an arbitrary temperature. Gorter and his collaborators did not use the anisotopy energy in the form (1·6) and their derivation of the critical hyperbola is not easy to see. If one uses (1·6), it is an easy task to obtain it, and indeed for an arbitrary temperature, not merely for absolute zero. This we will do in § 3.

As pointed out by Gorter and his collaborators, the anisotropy of the *g*-factor is really important. The *g*-values observed by Itoh and his collaborators[10] by paramagnetic resonance in CuCl₂·2H₂O at room temperature as follows :

$$g_a = 2.195, \quad g_b = 2.075, \quad g_c = 2.260.$$

These are not very much different from those obtained by the measurement of the susceptibility by van den Handel, Gijsman and Poulis[11] at low temperatures which are

$$g_a = 2.20, \quad g_b = 2.23, \quad g_c = 2.24.$$

If, therefore, we wish to employ an isotropic g equal to 2 in our theory, we have to replace the external field $\boldsymbol{H}$ appearing in theoretical formulas by $\boldsymbol{H}'$ defined by the following relation, which is numerically different from $\boldsymbol{H}$:

$$H_x' = (g_a/2) H_x, \quad H_y' = (g_b/2) H_y, \quad H_z' = (g_c/2) H_z. \tag{1·8}$$

At the same time, symbols for magnetization vectors, $\boldsymbol{M}^{\pm}$, $\boldsymbol{M} = \boldsymbol{M}^+ + \boldsymbol{M}^-$, $\boldsymbol{M}' = \boldsymbol{M}^+ - \boldsymbol{M}^-$, appearing in Yosida's paper are to be read $N\mu_B\langle \boldsymbol{S}^{\pm}\rangle$, $N\mu_B\langle \boldsymbol{S}\rangle$, $N\mu_B\langle \boldsymbol{S}'\rangle$ $(\boldsymbol{S} = \boldsymbol{S}^+ + \boldsymbol{S}^-$, $\boldsymbol{S}' = \boldsymbol{S}^+ - \boldsymbol{S}^-)$, where N is the total number of atoms and μ_B the Bohr magneton, and the parallel and perpendicular susceptibilities, $\chi_{\parallel}$ and $\chi_{\perp}$, are to be related to the measured susceptibilities in the following way :

$$\chi_{\parallel} = (2/g_a)^2 \chi_{aa}, \quad \chi_{\perp} = (2/g_b)^2 \chi_{ba} = (2/g_a)^2 \chi_{ab}. \tag{1·9}$$

Here χ_{ba}, for instance, is the susceptibility for the case where the applied field is parallel to the b-axis and the orientations of the spins parallel to the $+$ and $-$ a-axis. These facts may be evident in static cases, but might not be so in dynamic cases.

In order to see that, we write down the equation of motion :

$$i\hbar\dot{\boldsymbol{S}}^{\pm}=[\boldsymbol{S}^{\pm},\;\mathcal{H}],\qquad\qquad\qquad (1\cdot10)$$

where we take $(1\cdot5)$ for $\mathcal{H}$, if we are concerned with $\boldsymbol{S}^{+}$, and the corresponding expression when we are concerned with $\boldsymbol{S}^{-}$. We wish to treat $\langle\boldsymbol{S}^{\pm}\rangle$ as classical quantities. Then, using commutation relations for the components of angular momentum vector, we have

$$\dot{S}_{x}^{+}=a_{y}S_{z}^{+}\langle S_{y}^{-}\rangle-a_{z}S_{y}^{+}\langle S_{z}^{-}\rangle,\;\text{etc.}$$

For isotropic g this is nothing but (1.2) for vanishing external field. When there is an external field $\boldsymbol{H}$, the Hamiltonian contains the term $g_{a}\mu_{B}S_{x}^{+}H_{x}+g_{b}\mu_{B}S_{y}^{+}H_{y}+g_{c}\mu_{B}S_{z}^{+}H_{z}$ in the case of $\boldsymbol{S}^{+}$, and this indicates that we have only to replace $\boldsymbol{H}$ by $\boldsymbol{H}'$ in order to treat the spin magnetic moment as isotropic.

For the following we therefore consider the spin magnetic moment as isotropic and use $\boldsymbol{H}'$ in place of $\boldsymbol{H}$ but drop the prime for the sake of convenience. We also use the notation $\boldsymbol{M}^{+}$, $\boldsymbol{M}^{-}$, etc. and understand them as not being the true magnetization vectors but as $N\mu_{B}\langle\boldsymbol{S}^{+}\rangle$, etc. Also, we make use of the relation $(1\cdot9)$. The g-values we adopt in our numerical computations are those given by Itoh (1.c.), but they are not appreciably different from those given by van den Handel (1.c.). If the latter were used, our result for the resonance magnetic field in the ab-plane would have been a bit elongated in the direction of the b-axis owing to the difference between the values for g_{b}.

A few words might be added here. In the first place, we have tacitly assumed that the principal axes of the g-tensor coincide with the crystalline axes. In the case of $CuCl_{2}\cdot 2H_{2}O$, spins in the same c-plane are of the same sign but change their sign as we go from one plane to the next.[12] There are two sublattices in the same c-plane, corner atoms and face-centred atoms, and in general these two kinds of atoms should have their principal axes of g inclined symmetrically with respect to this c-plane. Actually, one of these axes coincides with the b-axis. Therefore, we are using the values of g averaged over these two sublattices. In the case of paramagnetic resonance and susceptibility measurements above the Néel point, this averaging is effected as

$$g_{a}^{2}=g_{\xi}^{2}\cos^{2}\alpha+g_{\zeta}^{2}\sin^{2}\alpha,\quad g_{c}^{2}=g_{\xi}^{2}\sin^{2}\alpha+g_{\zeta}^{2}\cos^{2}\alpha,$$

where α is the angle between one of the principal axes ξ and the a-axis. In our case, however, where we are concerned with phenomena below the Néel point, we have to take $g_{a}=g_{\xi}\cos^{2}\alpha+g_{\zeta}\sin^{2}\alpha$, etc., since we deal with magnetic moment directed along the a-axis, etc. These two averages are not the same if $\alpha\neq 0$. However, if the values of g_{ξ} and g_{ζ} are not much different from each other, both give almost the same result. The second is the question whether we can use the same value of g above and below the Néel point. In a recent paper Moriya and Yosida[13] calculated the anisotropy constants of $CuCl_{2}\cdot 2H_{2}O$ on the basis of dipole-dipole interaction and spin-orbit coupling connected with (super-) exchange effect. One sees in their paper that the ^{2}D ground state of Cu^{++} is split into five levels owing to the crystalline electric field, whose spacings are of the

order of $10000°K$. Since the exchange and anisotropy energies are thus very small compared with the level spacings, one is led to conclude that the g-values of individual atoms are the same below the Néel point as above.

§ 2. Anisotropy energy

We shall derive expression $(1 \cdot 6)$ for the anisotropy energy. We write the Weiss field acting on $M^{\pm}$ as

$$- (A_x M_x^{\mp} + B_x M_x^{\pm}, \; A_y M_y^{\mp} + B_y M_y^{\pm}, \; A_z M_z^{\mp} + B_z M_z^{\pm}). \qquad (2 \cdot 1)$$

The energy of the whole system per unit volume due to this field is

$$A_x M_x^{+} M_x^{-} + A_y M_y^{+} M_y^{-} + A_z M_z^{+} M_z^{-}$$
$$+ (1/2) (B_x M_x^{+2} + B_y M_y^{+2} + B_z M_z^{+2} + B_x M_x^{-2} + B_y M_y^{-2} + B_z M_z^{-2}). \qquad (2.2)$$

We shall now write

$$A_x = A, \; A_y = A + \mathit{\Delta} A_y, \; A_z = A + \mathit{\Delta} A_z,$$
$$B_x = B, \; B_y = B + \mathit{\Delta} B_y, \; B_z = B + \mathit{\Delta} B_z, \qquad (2 \cdot 3)$$

and consider the quantities attached with $\mathit{\Delta}$ to be small. Then the terms in (2.2) related with these quantities are as follows:

$$(1/2) (- \mathit{\Delta} A_y + \mathit{\Delta} B_y) (M_y^{+2} + M_y^{-2}) + (1/2) (- \mathit{\Delta} A_z + \mathit{\Delta} B_z) (M_z^{+2} + M_z^{-2}), \qquad (2 \cdot 4)$$

where we have used the relation

$$2 M_y^{+} M_y^{-} = (M_y^{+} + M_y^{-})^{2} - M_y^{+2} - M_y^{-2}, \text{ etc.,}$$

and have neglected $(M_y^{+} + M_y^{-})^{2}$, etc., since M^{+} and M^{-} are almost antiparallel. The anisotropy field (1.7) follows directly from (2.4) if we put

$$K_1 = M_0^2 (- \mathit{\Delta} A_y + \mathit{\Delta} B_y), \quad K_2 = M_0^2 (- \mathit{\Delta} A_z + \mathit{\Delta} B_z). \qquad (2 \cdot 5)$$

Here we mean by M_0 the magnitude of M^{+} and M^{-} corresponding to zero external magnetic field. In the presence of external field, these vectors are almost antiparallel, so that we can put $M_0 = (M^{+} + M^{-})/2$. If further we introduce their direction cosines, α_{+}, $\beta_{+}, \gamma_{+}, \alpha_{-}, \beta_{-}, \gamma_{-}$, then we can replace $M_y^{+2} + M_y^{-2}$, etc. by $M_0^2 (\beta_{+}^{2} + \beta_{-}^{2})$, etc. in (2.4), because

$$(M_y^{+2} + M_y^{-2}) - M_0^2 (\beta_{+}^{2} + \beta_{-}^{2})$$
$$= (M^{+2} - M_0^2) \beta_{+}^{2} + (M^{-2} - M_0^2) \beta_{-}^{2}$$
$$= (1/2) \; M_0 (M^{+} - M_0) (\beta_{+}^{2} - \beta_{-}^{2}) + 1/4 (M^{+} - M^{-})^{2} (\beta_{+}^{2} + \beta_{-}^{2}),$$

and the last two terms are of the second order of $M^{+} - M^{-}$ and can be neglected. Thus we have $(1 \cdot 6)$.

We shall Further define the common axis of the two vectors as

$$\alpha \propto (\alpha_{+} - \alpha_{-}), \; \beta \propto (\beta_{+} - \beta_{-}), \; \gamma \propto (\gamma_{+} - \gamma_{-}),$$

$$\alpha^2 + \beta^2 + \gamma^2 = 1. \tag{2.6}$$

Then the anisotropy energy can be written more simply as

$$K_1 \beta^2 + K_2 \gamma^2. \tag{2.7}$$

This will be used in deriving the critical hyperbola in the next section.

§ 3. Critical hyperbola

Néel (1.c.) predicted in 1936 a fact that, in the case of uniaxial anisotroy, there is a critical field such that, for the values of the field applied along the easy axis higher than this critical value, the axis of the antiparallel spins becomes perpendicular to the easy axis. This is because the perpendicular susceptibility is higher than the parallel susceptibility and the difference in the decreases of free energy due to perpendicular and parallel fields overcomes the difference in anisotropy energies for these two cases when the field strength exceeds the critical value.

As referred to in § 1, this prediction was generalized to the case of orthorhombic symmetry by Gorter and his co-workers. Experimentally, this has been confirmed by them in $CuCl_2 \cdot 2H_2O$. However, Gorter's theory is confined to absolute zero. The following theory is applicable to arbitrary temperature. It will be assumed that we have well-defined thermal average of the resultant spin on each sublattice, but we do not make use of the Weiss approximation.

The free energy may be expressed as the sum of three terms : the decrease of free energy due to the component of the applied field parallel to the spin axis, α, β, γ, that due to perpendicular component, and the anisotropy energy (2.7). We denote the direction cosines of the applied field as $\alpha_H, \beta_H, \gamma_H$. We then have the free energy :

$$-(1/2)\ \chi_{||}\ H^2(\alpha\alpha_H + \beta\beta_H + \gamma\gamma_H)^2 - 1/2\ \chi_{\perp}\ H^2[1 - (\alpha\alpha_H + \beta\beta_H + \gamma\gamma_H)^2]$$
$$+ K_1\beta^2 + K_2\gamma^2. \tag{3.1}$$

We have to find the minimum of this expression with respect to α, β, γ under a supplimentary condition $\alpha^2 + \beta^2 + \gamma^2 = 1$. This is a typical eigen-value problem. The secular equation is

$$\begin{vmatrix} \alpha_H^2 - \lambda & \alpha_H\beta_H & \alpha_H\gamma_H \\ \alpha_H\beta_H & \beta_H^2 + x_1 - \lambda & \beta_H\gamma_H \\ \alpha_H\gamma_H & \beta_H\gamma_H & \gamma_H^2 + x_2 - \lambda \end{vmatrix} = 0, \tag{3.2}$$

where

$$x_1 = \frac{2K_1}{(\chi_{\perp} - \chi_{||})H^2}, \qquad x_2 = \frac{2K_2}{(\chi_{\perp} - \chi_{||})H^2}. \tag{3.3}$$

The secular equation can be written as

$$\lambda^3 - (1 + x_1 + x_2)\lambda^2 + [(x_1 + x_2)\alpha_H^2 + x_2\beta_H^2 + x_1\gamma_H^2 + x_1 x_2]\lambda - x_1 x_2 \alpha_H^2 = 0. \tag{3.4}$$

There are three orthogonal directions of α, β, γ corresponding to the three roots of (3·4). We shall consider here three special cases.

1) The field is in the xz-plane, so that $\beta_H = 0$, $\gamma_H^2 = 1 - \alpha_H^2$. In this case (3·4) gives three roots :

$$\lambda = x_1, \quad \lambda = (1/2)\left[(1 + x_2) \pm \sqrt{(1 + x_2)^2 - 4x_2 \alpha_H^2}\right]. \tag{3·5}$$

The smallest value of λ corresponds to the stable state, for it represents the smallest value fo the free energy. Since the easy axis is in the x-direction and the most difficult axis in the z-direction, we assume $K_2 > K_1$ and consequently $x_2 \cdot x_1$. It is convenient to consider the roots as functions of α_H^2. When α_H^2 varies from 1 to 0, one of the roots, represented by the second formula of (3·5) with $+$ sign, goes up, another with $-$ sign goes down, and the remaining one, given by the first formula of (3·5) remains constant. These behaviours are shown schematically in Fig. 1.
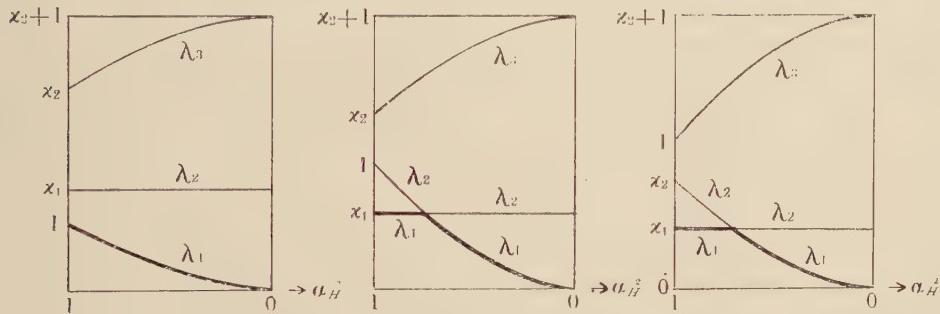

Fig. 1. Schematic representation of the behaviours of the roots λ_1, λ_2, λ_3. The lowest root λ_1 is realized.

In the cases (b) and (c) of Fig. 1, a crossing of two of the roots occurs. This crossing point is given for both cases by

$$\alpha^2 = x_1 (1 + x_2 - x_1) / x_2.$$

This can be written, using (3·3) and writing $H^2 \alpha_H^2 = H_x^2$, $H^2 = H_x^2 + H_z^2$, as

$$\frac{H_x^2}{2K_1} - \frac{H_z^2}{2(K_2 - K_1)} = \frac{1}{\chi_\perp - \chi_\parallel}, \tag{3·6}$$

which is nothing but the critical hyperbola of Gorter and Haantjes. In order to see that, we determine the ratio $\alpha : \beta : \gamma$ from the matrix of (3·2), which, in the present case, is

$$\begin{vmatrix} \alpha_H^2 - \lambda & 0 & \alpha_H \gamma_H \\ 0 & x_1 - \lambda & 0 \\ \alpha_H \gamma_H & 0 & \gamma_H^2 + x_2 - \lambda \end{vmatrix}.$$

If λ is equal to x_1, we have $\alpha : \beta : \gamma = 0 : 1 : 0$, so that the spins are directed along the y-axis. If, on the other hand, $\lambda \neq \kappa_1$, we have $\beta = 0$ and

$$\alpha : \gamma = \alpha_H \gamma_H : \lambda - \alpha_H^2. \tag{3·7}$$

This relation can be written in another form, using the angles ψ, θ defined in Fig. 2 (a) and substituting for λ the second minus formula of $(3\cdot5)$, as

$$\tan\ 2\psi = \frac{\sin\ 2\theta}{\cos\ 2\theta - 1/\varkappa_2}. \tag{3\cdot8}$$

This is the relation already derived by Néel (for absolute zero) and Yosida (for arbitrary temperatures). Geometrical representation of this relation is shown in Fig. 2. It may be seen that the spin axis is in the neighbourhood of the x-axis when θ is small or near to $\pi/2$ or H is small. Therefore the critical hyperbola derived above defines the point of the flop-over of the spin system from the xz-plane (the neighbourhood of the x-axis) to the y-axis.
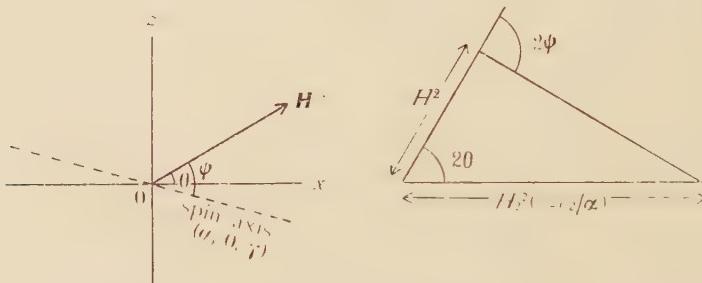

Fig. 2 (a)   Left : The definition of angles θ and ψ.
        (b)   Right : Relation between them.

2)   The field is in the xy-plane.   Similar arguments apply in this case.   We have only to exchange $\varkappa_1$ and $\varkappa_2$.   In particular, we obtain an ellipse

$$\frac{H_x^2}{2K_2} + \frac{H_y^2}{2(K_2 - K_1)} = \frac{1}{\chi_\perp - \chi_\parallel},$$

but this has no real meaning, since one gets it from the crossing of the two higher roots, of which one is $\lambda = \varkappa_{\cdots}$.   From mathematical point of view, it is interesting to notice that the critical hyperbola cuts the x-axis at the focal points of the above ellipse and vice versa. The only important relation is that which corresponds to $(3\cdot8)$.   It is

$$\tan\ 2\psi = \frac{\sin\ 2\theta}{\cos\ 2\theta - 1/\varkappa_1}. \tag{3\cdot9}$$

3)   The field is in the yz-plane.   The spin system orients always along the x-axis. There is nothing else to be mentioned particularly except that we obtain an imaginary quadratic curve.

§ 4   Theory of resonance

Theory of the resonance in a crystal of orthorhombic symmetry has been worked out by Yosida as an extension of the writer's theory of the case of uniaxial symmetry.   There is however, a peculiar resonance phenomenon at the critical field strength which has been

overlooked by Yoisida. This will be discussed in the next section, and in this section a summarizing review of Yosida's treatment with supplementary formulas will be given.

The problem is simply to solve the equation of motion (1·2) with the anisotropy field (1·7) in the neighbourhood of the static solution. It is convenient to add and subtract the equations for M^+ and M^- and to use the notations M and M' for the sum and difference of these vectors. In the right-hand sides of the equations for M and M', the latter can be replaced by a vector $2M_0(\alpha, \beta, \gamma)$ and the components of the former can be substituted, for instance in the case where the field is applied in the xy-plane, by

$$M_x = \chi_{||}H_x\cos(\psi-\theta) + \chi_\perp H_y\sin(\psi-\theta),$$

$$M_y = -\chi_{||}H_x\sin(\psi-\theta) + \chi_\perp H_y\cos(\psi-\theta).$$

Furthermore, we neglect the anisotry field when it is compared with the exchange field, and also neglect $\chi_\perp$ when it is compared with 1. In this way, Yosida obtained the following equation for the frequency of eigen-oscillation :

$$\left(\frac{\omega}{\gamma}\right)^4 - \left(\frac{\omega}{\gamma}\right)^2 \left[H^2(\alpha^2\cos^2\psi + 1) + c_2 + c_1(\cos^2(\psi-\theta) - 2\sin^2(\psi-\theta))\right]$$

$$+ H^4\alpha^2\cos^2\psi$$

$$- H^2[c_1(\alpha\cos^2\psi\cos2(\psi-\theta) + \alpha\sin\theta\cos\psi\sin(\psi-\theta) + \cos\theta\sin\psi\sin(\psi-\theta))$$

$$+ c_2(\alpha\cos^2\psi - \sin^2\psi)]$$

$$+ c_1\cos2(\psi-\theta)(c_2 - c_1\sin^2(\psi-\theta)) = 0. \tag{4·1}$$

This is valid when the external field is applied in the xy-plane. The notation here used is as follows :

$$c_1 = 2K_1/\chi_\perp, \quad c_2 = 2K_2/\chi_\perp, \quad \alpha = 1 - \chi_{||}/\chi_\perp. \tag{4·2}$$

We made use of the fact that $\chi_\perp = 1/\Delta$ in the Weiss approximation. In all our formulas, including that for the critical hyperbola, no parameter other than c_1, c_2 and α appears, and we shall determine these parameters by comparing our theory with a part of the experimental data. This way of determining the parameters eliminates to a certain measure the approximate character of the Weiss field.

When the field is applied in the xz-plane, and is inside the critical hyperbola, we obtain the same equation as (4·1) with c_1 and c_2 exchanged. When the field is outside it, the equation of small oscillation is simpler, since the spins direct along the y-axis, and the following equation can be obtained :

$$\left(\frac{\omega}{\gamma}\right)^4 - \left(\frac{\omega}{\gamma}\right)^2 (H^2 + c_2 - 2c_1) + (c_2 - c_1)H_x^2 - c_1H_z^2 - c_1(c_2 - c_1) = 0.$$

This can be written in a more elegant form :

$$\frac{H_x^2}{c_1 + (\omega/\gamma)^2} - \frac{H_z^2}{(c_2 - c_1) - (\omega/\gamma)^2} = 1. \tag{4·3}$$

This we will call *resonance hyperbola.*

When the external field is applied along the x-axis, we have from $(4 \cdot 1)$, putting $\theta = \phi' = 0$,

$$\left(\frac{\omega}{\gamma}\right)^4 - \left(\frac{\omega}{\gamma}\right)^2 [H^2(a^2+1)+c_2+c_1] + H^4 a^2 - H^2(c_1+c_2)a + c_1 c_2 = 0. \qquad (4 \cdot 4)$$

Therefore

$$\left(\frac{\omega}{\gamma}\right)^2 = \frac{1}{2}[H^2(a^2+1)+c_2+c_1$$

$$\pm \{H^4(1-a^2)^2 + 2H^2(1+a)^2(c_1+c_2) + (c_2-c_1)^2\}^{1/2}]. \qquad (4 \cdot 5)$$

This is valid when H is smaller than $H_c = (c_1/a)^{1/2}$, the value of the critical field strength on the x-axis. When H is greater than H_c, we have from $(4 \cdot 3)$, putting $H_z = 0$,

$$\left(\frac{\omega}{\gamma}\right)^2 = H^2 - c_1, \quad \left(\frac{\omega}{\gamma}\right)^2 = c_2 - c_1. \qquad (4 \cdot 6)$$

The mode of oscillation corresponding to the first of $(4 \cdot 6)$ is such that M^+ and M^- (which are almost parallel to $+$ and $-y$-axis and are tilted a little towards the x-axis by the action of the applied field) are rigidly connected with each other and make a rotatinal oscillation with the z-axis as the rotational axis and, with a phase difference of $\pi/2$, another rotational oscillation with the y-axis as the rotational axis. This mode is therefore excited by the component of the oscillating magnetic field in the yz-plane, the component along the z-direction being most effective, especially at low temperatures. The mode corresponding to the second of $(4 \cdot 6)$ is such that the mentioned two vectors change the angle between them, symmetrically with respect to the xz-plane, and with a phase difference of $\pi/2$ make a rotational oscillation with the x-axis as the rotational axis. This mode is excitable active only for the x-component of the oscillating field.

The modes corresponding to $(4 \cdot 5)$ are such that M^+ oscillates with a larger amplitude than M^- for the $+$ sign of $(4 \cdot 5)$ and the reverse holds for the $-$ sign. These are with an oscillating field perpendicular to the x-axis.

When the static external field is along the y-axis, we have from $(4 \cdot 1)$, putting $\theta = \phi' = \pi/2$,

$$\left(\frac{\omega}{\gamma}\right)^4 - \left(\frac{\omega}{\gamma}\right)^2 (H^2+c_2+c_1) + H^2 c_2 + c_1 c_2 = 0,$$

so that

$$\left(\frac{\omega}{\gamma}\right)^2 = H^2 + c_1, \quad \left(\frac{\omega}{\gamma}\right)^2 = c_2. \qquad (4 \cdot 7)$$

The corresponding modes are similar to those of (4.6), x and y being interchanged.

Finally, the eigen-frequency in the yz-plane is given by the equation

$$\frac{H_y^2}{(\omega/\gamma)^2 - c_1} + \frac{H_z^2}{(\omega/\gamma)^2 - c_2} = 1. \qquad (4 \cdot 8)$$

§ 5. Resonance at the critical field

When the external static field is applied along the a-axis and its strength crosses just over the critical field strength H_c from the lower side, the spin direction turns suddenly from the a-axis to the b-axis along the ab-plane. (In actual crystals this turning may occur in a certain small interval of H, depending on the extent to which the crystals possess internal irregularities.) At each orientation the spin system has the corresponding eigen-oscillation and eigen-frequency. When the eigen-frequency at some orientation coincides with the frequency of the applied oscillating magnetic field, resonance takes place at H_c. We shall investigate the condition for this in the following lines. We shall confine ourselves to the case of static field applied along the a-axis ; similar resonance phenomenon may occur when the field is on the critical hyperbola, but there seems to be a mathematical difficulty in treating this case, since we have to determine the three-dimensional path along which the spin vectors move.

We can make use of equation (4·1). We put $\theta=0$ (θ is the angle between the a-axis and the direction of the applied field) and $H^2=H_c^2=c_1/\alpha$. Then we have

$$\left(\frac{\omega}{\gamma}\right)^4 - \left(\frac{\omega}{\gamma}\right)^2 \left[(c_2-c_1) + c_1(\alpha^{-1}+2+\alpha) - c_1(3+\alpha)\sin^2\psi\right]$$

$$+ (c_2-c_1)c_1(\alpha^{-1}-1)\sin^2\psi=0. \tag{5·1}$$

For $\psi=0$ we have $(\omega/\gamma)^2=0$ and $(\omega/\gamma)^2=(c_2-c_1)+c_1(\alpha^{-1}+2+\alpha)$. With increasing ψ the coefficient of $-(\omega/\gamma)^2$ in (5·1) decreases and the last term increases, so that one of the roots increases and the other decreases. At $\psi=\pi/2$, the roots are $(\omega/\gamma)^2=c_2-c_1$ and $(\omega/\gamma)^2=c_1(\alpha^{-1}-1)$. This behaviour of the roots is schematically shown in Fig. 3.
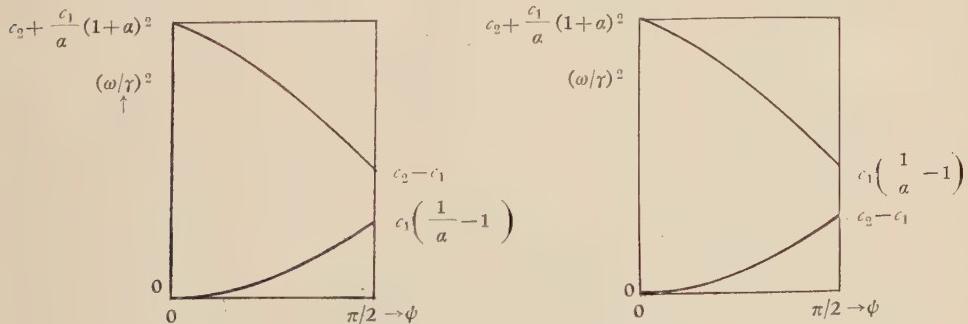

Fig. 3. Schematic representation of the roots of equation (1.1).

There are two cases, (a) and (b), shown in Fig. 3.

a) $c_2-c_1>c_1(\alpha^{-1}-1)$. This is valid at low temperatures, since α tends to 1 as the temperature goes down to zero. c_1 and c_2 are proportional to M_0^2 so that they are almost constant at lower temperatures but decreases rapidly to zero when the temperature approaches the Néel point. c_1/α is known, in the case of CuCl$_2$·2H$_2$O, to be an increasing function of temperature. Therefore the above inequality breaks down at a certain tem-

perature ($3.69°K$ for $CuCl_2 \cdot 2H_2O$, the Néel temperature beidg $4.33°K$). If the circular frequency of the applied oscillating field, ω_0, satisfies the relation, either

$$(\omega_0/\gamma)^2 > c_2 + c_1(\alpha^{-1} + 2 + \alpha)$$

or

$$c_2 - c_1 > (\omega_0/\gamma)^2 > c_1(\alpha^{-1} - 1),$$

no resonance can take place during the turning of the spins, as one may judge from Fig. 3(a). This is the case for $T < 2.25°K$ in Ubbink's experiment with $CuCl_2 \cdot 2H_2O$, using 3 cm wave. In other cases resonance can occur.

  b) $c_1(\alpha^{-1} - 1) > c_2 - c_1$, that is, $c_1 \alpha^{-1} = H_c^2 > c_2$. This is valid for higher temperatures. In order to have a resonance at the critical field, either

$$c_2 - c_1 > (\omega_0/\gamma)^2$$

or

$$c_2 + c_1(\alpha^{-1} + 2 + \alpha) > (\omega_0/\gamma)^2 > c_1(\alpha^{-1} - 1)$$

must be satisfied. When the temperature rises, the first condition ceases to be valid at a
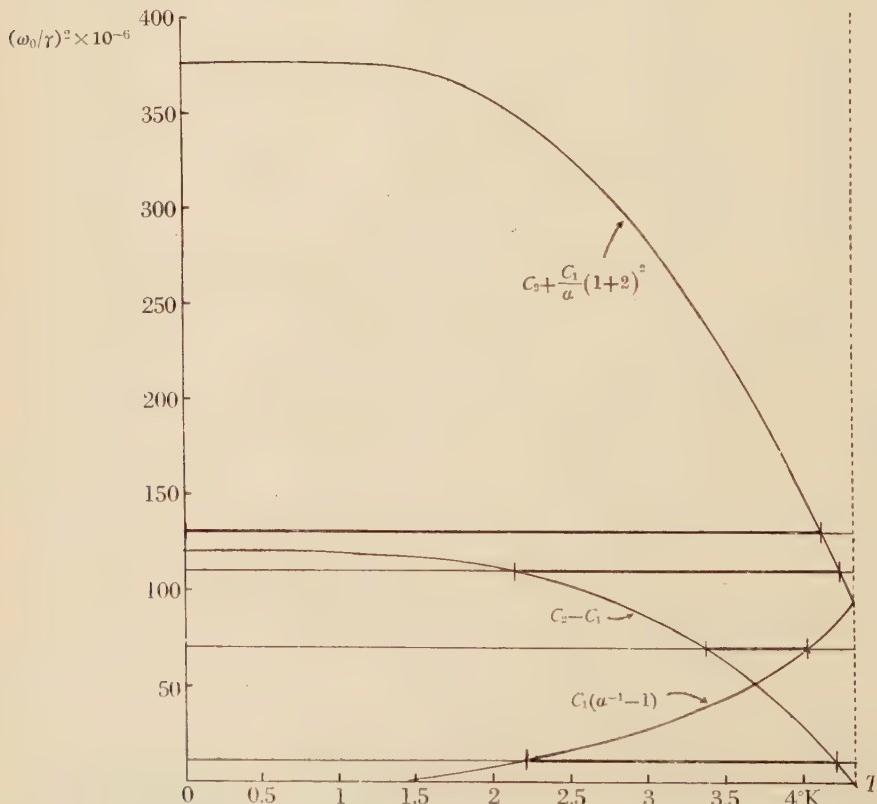
Fig. 4. Diagram for obtaining the range of temperature in which resonance takes place at $H = H_c$. Thick parts of horizontal lines show them; the heights of these horizontal lines are different assumed values of $(\omega_0/\gamma) \times^2 10^{-6}$. This diagram is applicable numerically to $CuCl_2 \cdot 2H_2O$.

certain temperature, since c_1 and c_2 are both decreasing functions of temperature. In Ubbink's experiment with 3 cm wave the second condition is always not satisfied.

It is more convenient to plot the three functions $c_2 + c_1(a^{-1} + 2 + a)$, $c_2 - c_1$, $c_1(a^{-1} - 1)$ versus T and to see in what region of temperature resonance takes place at the critical field. Horizontal lines in Fig. 4 show different assumed values of $(\omega_0/\gamma)^2$ by their height and the thick parts of them indicate the range of temperature in which resonance takes place at $H = H_c$. It is interesting to notice that at the point where a horizontal curve meets the curve of $c_2 - c_1$ the second resonance condition of (4·6) is satisfied, so that resonance takes place at any field strength higher than H_c, resulting in an absorption curve of the shape flat for $H > H_c$ and cut off for $H < H_c$. Such an absorption curve has actually been observed by Ubbink, and our theory gives explanation for that.

§ 6. Comparison between theory and experiment

We shall now compare our theoretical results with Ubbink's experiment of resonance absorption in CuCl₂·2H₂O with 3 cm wave. We look for the resonance value of H corresponding thus to $(\omega/\gamma)^2 = 11.3 \times 10^6$. There are three parameters in our theory, c_1, c_2 and a. c_1 and c_2 are theoretically proportional to M_0^2 and so constant at low temperatures but decrease as the temperature approaches the Néel point and vanish there. The ratio c_2/c_1 should, however, be constant, independent of temperature, although in actuality it might depend slightly upon temperature owing to lattice expansion. We shall use the value $c_2/c_1 = 3.3$ which was found by the method of trial and error. $a = 1 - \chi_{11}/\chi_{\perp}$ can be found from the measurements of susceptibilities, using (1·9), but $\chi_{\perp}$ depends on whether we use $\chi_{a'}$ or χ_{ba} and also depends on whether we use high field value or low field value of them.[11] The value of a actually adopted is shown in Fig. 5, together with experimental points. A slight change in the choice of the value of a affects the results rather drastically, so that a number of trials have been made before deciding this curve. We are finally left the determination of c_1 as a function of temperature. This can be done by comparing theory with the experiment of resonance in the case where the external field is applied along the a-axis, as will be mentioned below. The determination of c_1 will give M_0, the spontaneous magnetization of each sublattice, as a function of temperature. This can be compared with the curve of M_0 determined from the proton resonance measurements of Poulis,[12] that determined from the specific heat measurements of Friedberg,[14] or that determined theoretically using Weiss approximation. This comparison is already made in Ubbink's thesis (his Fig. 3, p. 29), and the curve determined from the proton resonance measurements lies the highest, that from specific heat measurements next, that from c_1 thirdly, and that from Weiss approximation the lowest. Ubbink has indeed already made comparison between his experiment and Yosida's theoretical result in the case where the field is applied along the a-axis and also his own theory at absolute zero for the cases where the field is applied in the ac-plane and ab-plane.

a) The field is applied along the a-axis. In this case, our theoretical considerations predict resonance at H_c in a certain interval of temperature, which is actually between 2.25°K and 4.22₅°K. (The Néel point is 4.33°K according to Ubbink's paper, but

specific heat measurements give 4.316°K, according to Prof. Gorter's private communica-
tion; this, however, does not matter in the present comparison, because the parts of the
curves in Fig. 6 above, say, 4.25°K have been drawn without numerical calculations.)
The resonance along the H_c-curve was suggested by Yosida without any sound theoretical consi-
derations, and Ubbink made use of this suggestion in explaining his results. The upper
ascending curve in Fig. 6 indicates Ubbink's values of H_c obtaind in this way. Ubbink
noticed furthermore that the resonance condition given by the second equation of (4·6) is
satisfied just at the point where the resonance at H_c vanishes, and therefore that at this
point (4.22, K) a semi-infinitely broad resonance absorption curve should be obtained, just
as his experiment shows. These were actually correct.

From the H_c-curve we can obtain the values of c_1 in the corresponding temperature
range. Below this range, the value of H given by the first resonance condition of (4·6)
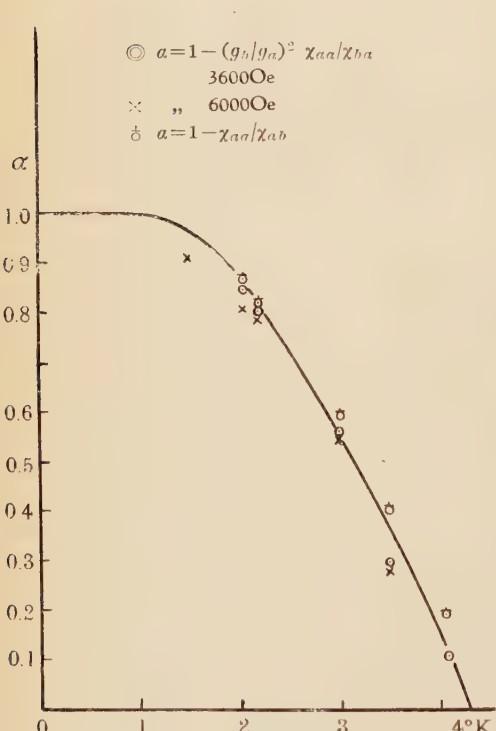

Fig. 5. Curve for a.
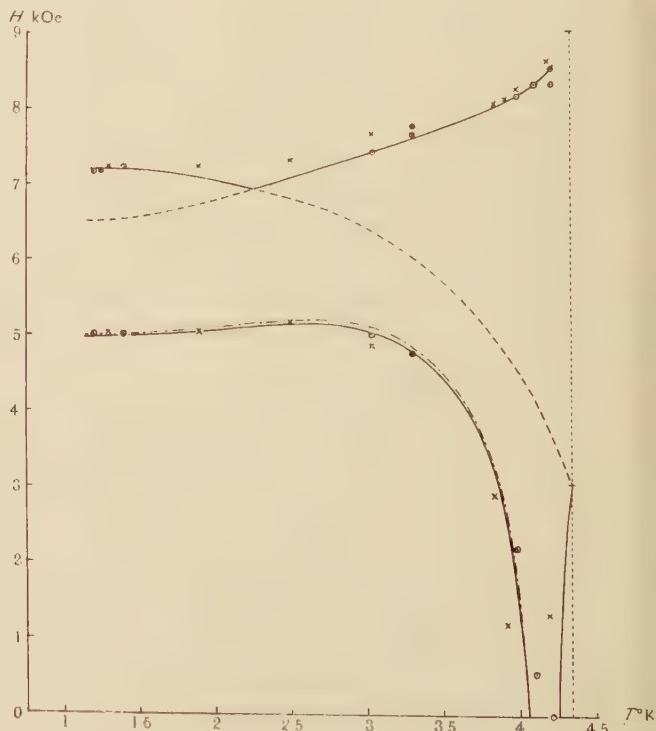
Fig. 6. Resonance field strengths when the
field is applied along the a-axis.

becomes higher than H_c, so that this condition becomes valid and at the same time the
condition of resonance at H_c becomes to fail. Therefore we obtain the resonance along the full
curves given in the upper side of Fig. 6. At the lowest temperature, 1.2°K, the value
of c_1 can be determined from this curve and from the value of $a(=0.995)$ given by the curve
of Fig. 5; this value of a was actually so chosen as to obtain agreement with the lower resonance

$C_1 \times 10^{-6}$
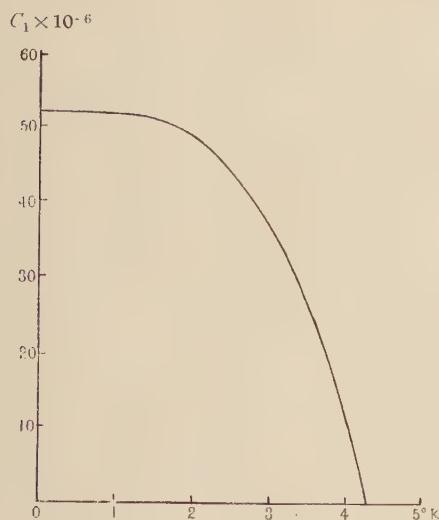
Fig. 7   Curve for c_1

curve of Fig. 6. Between 1.2°K and a higher temperature, say 2.5°K, the value of c_1 can be determined by interpolation. In this way we obtained the whole upper two curves, full and broken. Then, taking $c_2/c_1 = 3.3$ we obtain the lower resonance curves, using (4·4). (With $c_2/c_1 = 3.5$, we would have obtained the chain curve indicated in the same figure.)

Almost the same figure has been obtained by Ubbink through independent numerical calculations (Fig. 7, p. 66 in his thesis).

We see that the agreement is surprisingly good, surprisingly because our theory is based on Weiss approximation.

b)   The field is applied in the ac-plane. We can now calculate straightforwardly the resonance curves in the ac-plane. The results are shown in Fig. 8.

Lower curves in Fig. 8 for 1.3°K, 3.88°K, and 3.925°K were calculated from (4·1),
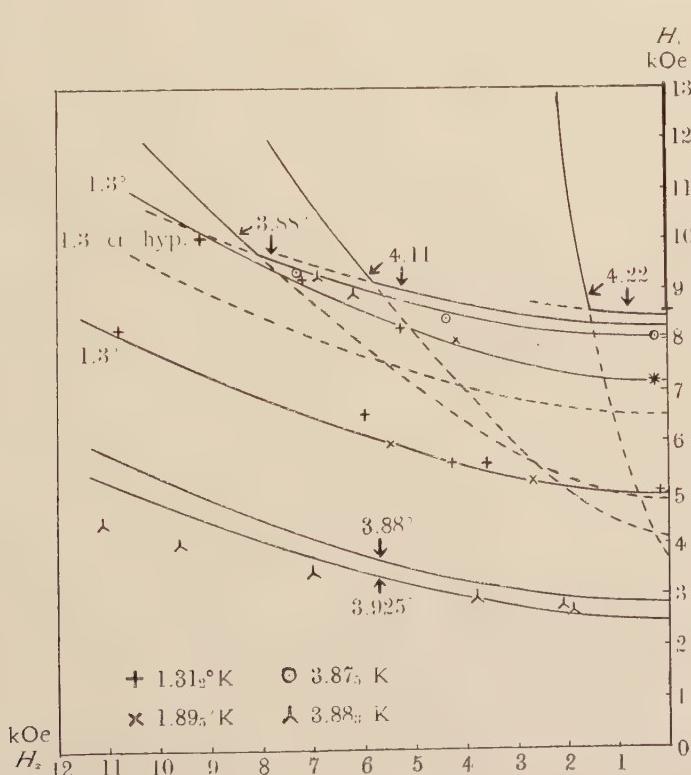
Fig. 8.   Resonance curves in the ac-plane.
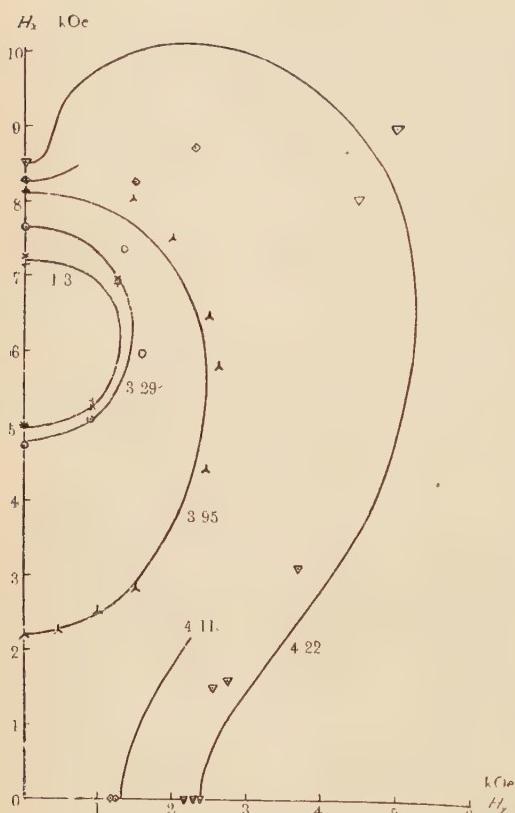
Fig. 9. Resonance curves in the ab-plane. Experimental points : + 1.23$_5$°K and 1.27°K, × 1.39°K and 1.43°K, ◎ 3.29$_5$°K, ⅄ 3.98$_3$°K, ◈ 4.11°K and 4.15°K, ▽ 4.22°K.

with c_1 and c_2 interchanged, and with (3·8). The curve calculated for 1.3°K is in excellent agreement with experimental ponits. The curve for 3.88°K lies higher than the experimental points, but as one sees from Fig. 6 a small change in temperature causes a great change in the resonance value of H, so we took arbitrarily another temperature 3.925°K to compare with experimental points for 3.88°K.

The upper curve for 1.3°K is the resonance hyperbola (4·3), and it lies higher than the critical hyperbola. For 3.88°K the critical hyperbola lies higher than the resonance hyperbola for H_z smaller than about 8 kOe, but at this value of H_z these hyperbolas cross with each other. We may expect that resonance takes place along the critical hyperbola for H_z smaller than 8 kOe, because this is true for both end points, and the experimental points lie actually on this critical hyperbola. Beyond this value of H_z, we should have resonance along the resonance hyperbola, but there is no experimental point.

With increasing temperature the cross point of the two hyperbolas comes nearer to the x-axis as one sees in the figure, and finally at 4.22°K the cross point comes on the x-axis, the resonance hyperbola coinciding with the x-axis. At this temperature we obtain resonance for any value of H greater than H_c on the x-axis. At still higher temperatures, we should have no resonance.

c) The field is applied in the ab-plane. In this case there is no such remarkable features as those we saw in the case of the ac-plane. Calculated curves fit well with experiment, as one sees in Fig. 9. except that experimental points lie a little outside the calculated curves at the upper right side of Fig. 9.

Line Width and Polarization Effect Although detailed observations of the line width have not been made, Ubbink states that the absorption lines broaden as the direction of the applied field is turned from the a-axis in both the ac- and ab-planes. Apart from the fact that at higher temperatures the peak position for lower field strengths is sensitively affected by a slight change in temperature, we might mention two reasons for this broadening. One is that, in the case of ac-plane, the resonance curves are increasing functions of H_z so that a straight line passing through the origin cuts these curves obliquely. With decreas-

ing cutting angle the absorption curve may broaden. The same thing can be said for the lower part of Fig. 9 for the ab-case. The other reason is as follows and may probably be more important, especially in the ab-case, and in particular when H is not much different from H_c in this case. As the direction of the field is turned from the a-axis the spin direction deviates from this axis, too. (See fig. 2.) In real crystals there may be internal strains and other kinds of imperfections, so the spin system is subjected to the force due to them and its direction may be spread over a certain range of angle. We thus have to expect a broadening, and the line must be broader for the region where the direction of spins is affected more sensitively by the change of the field strength.

Some polarization effects are also reported by Ubbink. In the ab-case he uses two modes of cavity oscillation. In the first mode the direction of the oscillating magnetic field is inclined to the a-axis by $15°$, and in the second mode by $-75°$, both for 3 cm. With H parallel to the a-axis the resonance absorption at $H > H_c$ in the first mode is considerably higher than in the second mode, and that at $H < H_c$ in the first mode is considerably lower than in the second mode, at lower temperatures. The second of these statements is understood from our theory, since the oscillation below the critical field is excitable with oscillating field perpendicular to the a-axis (most excitable with oscillating field parallel to the c-axis, but that parallel to the b-axis is also efficient). The first statement is curious, because the oscillation above the critical field (the first equation of $(4 \cdot 6)$) is excitable only by the component of the oscillating field perpendicular to the a-axis, the c-component being again most effective while the b-component being almost inefficient. The writer suspects if there was not the c-component in Ubbink's experiment and that it was not stronger in the first cavity mode than in the second. It whould have been the dominant component for the excitation. At the highest temperature $(4.22°K)$ the absorption in the first mode was observed to be higher than that in the second mode over the whole resonance curve. Theoretical consideration is difficult except either on the a-axis or on the b-axis, but it predicts that the first mode is more efficint on the b-axis while the reverse is true on the a-axis. The latter is in contradiction with observation. The writer feels that it is necessary to have more precise data.

§ 7. **Further theoretical predictions**

Corresponding to four horizontal lines in Fig. 4 we have drawn three more sets of resonance curves with H parallel to the a-axis, the case of the lowest lying horizontal line being already shown in Fig. 6. Fig. 10 (a), (b), (c) show them, full parts of these curves representing the ranges where resonance occurs. In (a) and (b) one gets a perpendicular line starting from a point on the curve of H_c and going to infinity. These figures show what one would get with $CuCl_2 \cdot 2H_2O$ for 1.45 cm, 0.93 cm and 0.78 cm waves, respectively.

We can also plot ω/γ as a function of H for a given temperature. Fig. 11 (a), (b), (c) show this for $T = 0$, $3.03°K$, and $3.88°K$. The full parts of the perpendicular lines at $H = H_c$ show the ranges where resonance occurs at this field strength.
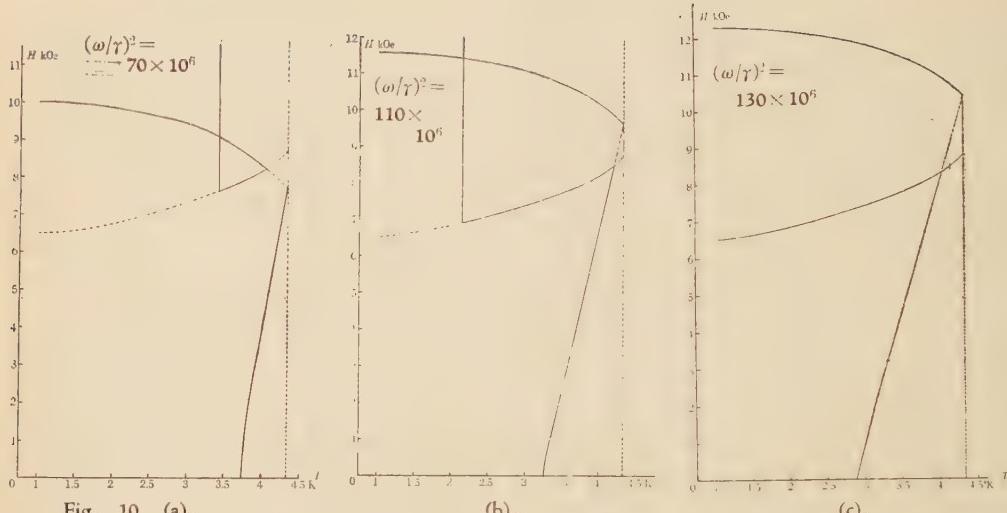
Fig. 10 (a), (b), (c).
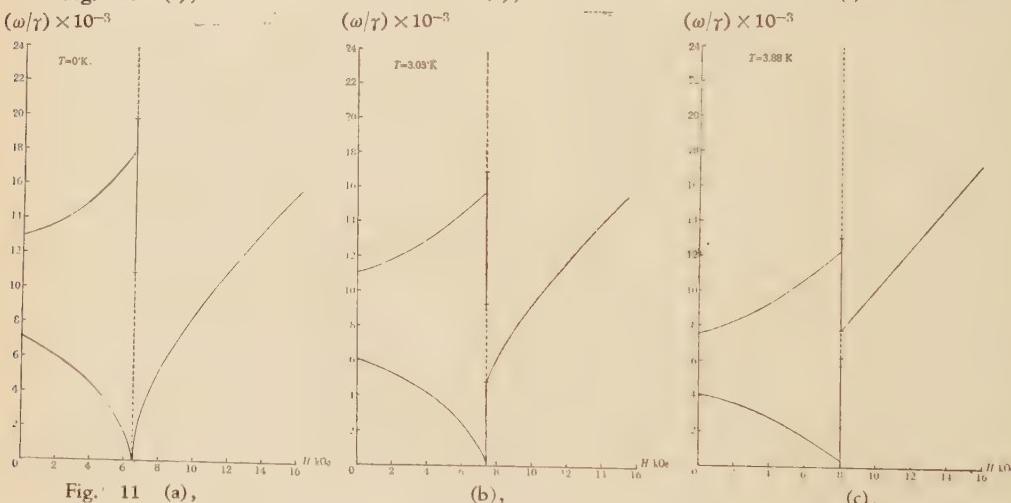
Fig. 11 (a), (b), (c).

§ 8. Discussion

The only point to be criticized with this theory might be that it is based upon the Weiss approximation, though this approximation is not fully made use of. In deriving the critical hyperbola we assumed that there are well-defined magnetization vectors of the two sublattices whose directions are almost antiparallel along the direction denoted as a, β, γ, and we further assumed that the susceptibility along this direction and that perpendicular to it are obtained experimentally through the relations (1·9). No further assumptions were made. In treating the resonance phenomena the equation of motion was written using the Weiss approximation, we made use of the relation $A = 1/\chi_{\perp}$, but otherwise relied upon using experimental data to determine unknown parameters, c_1, c_2, and a, c_2/c_1 being assumed to be constant. The way so followed must have got rid of certain weaknesses inherent in the Weiss approximation. There are a number of facts against the Weiss approxi-

mation:[*] the specific heat anomaly is not such that one would expect from this approximation, and the anomalous entropy change from absolute zero up to the Néel point is only 3.77 joules/deg. mol., according to Friedberg[14], as compared with $Rln\ 2 = 5.78$, the remaining entropy change bistributed above the Néel point. The spontaneous magnetization of each sublattice as a function of temperature measured by proton resonance and other methods is not the same as that derived from the Weiss approximation. Similar is the case with susceptibilities. However, our theoretical results explain the experimental results of resonance so well that it might be said that the way followed by the present theory can at least semi-empirically be applied safely to such a kind of resonance phenomena.

References

1) T. Nagamiya, Prog. Theor. Phys. **6**(1951), 432, 350, also F. Keffer and C. Kittel, Phys. Rev. **85** (1952), 329.
2) K. Yosida, Prog. Theor. Phys. **6** (1951), 691.
3) K. Yosida, Prog. Theor. Phys. **7** (1952), 25.
4) K. Yosida, Prog. Theor. Phys. **7** (1952), 425.
5) J. H. van Vleck, Phys. Rev. **52** (1937), 1178.
6) N. J. Poulis, J. van den Handel, J. Ubbink, J. A. Poulis and C. J. Gorter, Phys. Rev. **82** (1951), 552; J. Ubbink, J. A. Poulis, H. J. Gerritsen and C. J. Gorter, Physica **18** (1952), 361; J. Ubbink, Thesis (Leiden, 1953).
7) C. J. Gorter and J. Haantjes, Physica **18** (1952), 285.
8) J. Ubbink, Thesis, 1. c.; Physica **19** (1953), 9.
9) L. Néel, Ann. de Phys. 11e Ser. **5** (1936), 232.
10) J. Itoh, M. Fujimoto and H. Ibamoto, Phys. Rev. **83** (1951), 852.
11) J. van den Handel, H. M. Gijsman and N. J. Poulis, Physica **18** (1952), 862.
12) N. J. Poulis, Thesis (Leiden, 1952); N. J. Poulis and G. E. G. Hardeman, Physica **18** (1952), 201, 315.
13) T. Moriya and K. Yosida, Prog. Theor. Phys. **9** (1953), 663.
14) S. A. Friedberg, Physica **18** (1952), 714.
*) Prof. Gorter pointed out these in detail in the Confence on Theoretical Physics in Kyoto, September 1953.

Letters to the Editor

A Note on the Generalized Statistics

F. Ferrari* and C. C. Villi**

Istituto Nazionale di Fisica Nucleare,
Sezione di Padova, Italy

March 8, 1954

Method of field quantization more general than the usual ones in accordance with Bose-Einstein and Fermi-Dirac statistics have been recently investigated.[1] [2] The essential assumption underlying them is the validity of a generalized statistics which, as well as the Fermi statistics, is not necessarily deduced from quantum mechanics, but established by a generalized exclusion principle devised ad hoc and allowing a state to be occupied at most with d particles.

The average number of constituent systems in the j-th state, satisfying the conditions

$$\sum N_j = N \; ; \quad \sum N_j \varepsilon_j = E$$

is then given by

$$(1) \qquad f(\varepsilon) = \frac{\sum W N_j}{\sum W}$$

where the statistical weight W is subjected to the conditions

$$(2) \quad \begin{array}{l} W=1 \quad \text{for} \quad N_j = 0, 1, 2, \cdots\cdots d \quad (\text{any } j) \\ W=0 \quad \text{for} \quad N_j > d. \end{array}$$

From the point of view of quantum mechanics, conditions (2) are taken into account by letting the wave function Ψ of an assembly, obeying intermediate statistical rules, to be different from zero or zero depending on whether $0 \leq N_j \leq d$ or, respectively, $N_j > d$. This specification would indeed appear to arise difficulties concerning the symmetry properties of wave functions; they can be partially overcome by assuming a priori that, as the Bose-Einstein and the Fermi-Dirac statistics are related to the identical and to the alternating representation of the symmetric group, intermediate statistics are associated to the other representations of this group.[3]

It may be noted that both Okayama generalized statistics and Green's generalized method of field

* Department of Physics, Padua University, Italy.
** Department of Physics, Trieste University, Italy.

quantization involve an intermediate statistics characterized by the distribution function

$$(3) \quad f(\varepsilon) = \frac{1}{(\vartheta^\varepsilon/\lambda) - 1} - \frac{d+1}{(\vartheta^\varepsilon/\lambda)^{d+1} - 1}$$

where $\lambda = \exp(\eta(T)/kT)$ and $\vartheta = \exp(1/kT)$. From the d.f. (3), originally derived by Gentile[4] by means of the Bose method, it is easily seen that the two limiting Bose-Einstein and Fermi-Dirac cases are obtained on taking $d = \infty$ and, respectively, $d = 1$ for the maximum number of particles in any one non degenerate energy level.

It is immediate to prove that the function $W(\varepsilon)$, given by Okayama, reduces to the distribution function (3) on performing the summation of the d-th order sequences

$$\sum_{1}^{d} \lambda^\rho \vartheta^{-\varepsilon\rho} \quad (\varepsilon > \eta)$$

and

$$\sum_{1}^{\infty} (d - \rho + 1) \lambda^{1-\rho} \vartheta^{\varepsilon(\rho-1)}, \quad (\varepsilon < \eta).$$

The same result is also achieved either by solving Eq. (1) according to the Darwin-Fowler method or diagonalizing the energy momentum 4-vector of a half spin field, obeying Green's generalized rules of field quantization, and writing in matrix from $H = \sum_j N \varepsilon_j$ the total Hamiltonian of the assembly. In the latter case, being the quantum Zustandsumme

$$Z = \text{trace } (\lambda^N \vartheta^{-H})$$

and dropping off the condition that the number of particles in a state be 0 or 1, but at most d, one gets

$$Z = \sum_{N_j} \lambda^{\sum_j N_j} \vartheta^{-\sum_j N_j \varepsilon_j} = \prod_j \Omega_j$$

where

$$\Omega_j = \frac{1 - (\lambda \vartheta^{-\varepsilon_j})^{d+1}}{1 - \lambda \vartheta^{-\varepsilon_j}}$$

and hence the d.f. (3) is derived

$$f(\varepsilon) = \lambda \frac{\partial}{\partial \lambda} \log \Omega$$

Thus, Okayama and Green's attempts imply the validity of the Gentile statistics. The generalized exclusion principle, on which this statistics is based, requires that N satisfies a d-th order algebraic equation having as eigenvalues n_ρ all integer number from 0 to d, i.e.

$$\prod_{0}^{d} {}_\rho (N - \eta_\rho) = 0.$$

In particular, the Okayama result is found on substituting N in the above equation with the creation-destruction operator XX^*.

Following an Okayama suggestion, it may be interesting to inquire whether an intermediate statistics based on the d.f. (3), so far applied only to the study of Helium II,[4] could provide at least a useful statistical tool for the description of assemblies formed by strongly interacting particles. Of course, a crucial application of such a method is offered by an aggregate of particles like a heavy nucleus of mass number A.

Calculation of the excitation energy E by means of the d.f. (3), through series expansion of $[(\vartheta^\varepsilon/\lambda)-1]^{-1}$ and $[(\vartheta^\varepsilon/\lambda)^{d+1}-1]^{-1}$, leads again to a dependence on temperature according to the law $E=\gamma a T^2$, where $\gamma(d)=2d^{-1/3}(1+d)^{-1}$ and $a=(\pi^2/4)(A/\eta_0)$ is the usual coefficient for the Fermi ideal case $(d=1)$.

The corrections thus involved by an intermediate statistics for $d>1$ modify the theoretical previsions on the specific heat of nuclear matter and the nuclear level density in the same direction either of the Bardeen correlated model[5] or of models based on a hot Fermi distribution.[6]

This agreement might be understood on considering that either the reduction of the nucleon maximum momentum by a factor $(1/d)^{1/3}$ with respect to the Fermi gas model at $T=0$, expected according with the d.f. (3), or a smoothed down Fermi distribution, corresponding to $T\neq0$, allows a certain number of transitions, otherwise forbidden by the exclusion principle. However, the loosening of the Fermi restrictions, whatever may be the way that the exclusion principle is made less operative, does not provide an adequate theoretical outline from which to get a correct description of the behavior of assemblies formed by strongly interacting particles.

In fact, in order to put theoretical expectations in a somewhat good agreement with experimental data on π^+-meson production cross sections for nucleons against nuclei, it seems necessary to assume that nucleons inside the nucleus have a maximum momentum larger and not smaller than the maximum momentum corresponding to the Fermi gas model at $T=0$.[7]

This requirement may be accomplished with by means of a hot Fermi or Golberger-Chew distribution for nucleon momenta and not with an intermediate one, because the reduction of the maximum nucleon momentum, brought about by the Gentile statistics, has on the cross section for meson production a

bearing predominant respect that due to the limitation of the Pauli principle. This argument, however, probably is not too conclusive, because calculations of π-meson production cross sections are generally performed on the basis of free nucleon cross sections; this is obviously rather objectionable with reference to a model, like that outlined by the intermediate statistics, which claims to describe assemblies of strongly interacting particles.

Furthermore, it has to be pointed out that an intermediate statistics, based on the d.f. (3), does not give a correct account on the positive to negative π-meson ratio,[8] for it leads to a disfavoring of positives to negatives when the incident particle is a proton and vice versa when it is a neutron, i.e. the ratio π^+/π^- tends to unity when d tends to infinite.

The tentative generalization of statistics pursued by Lindsay[9] assigning, according to Brillouin, to each cell with d occupants a capacity $1-ad$, where a is a real positive parameter $0\leq a\leq1$, leads to the same results as those derived by the Gentile statistics. The number of ways of distributing all N particles among g_j cells according with the above capacity limitations is

$$w_j=a^{N_j}(g_j/a)!\{[(g_j/a)-N_j]!\}^{-1}.$$

The d.f. is then

$$f(\varepsilon)=\frac{1}{(\vartheta^\varepsilon/\lambda)+a}$$

which reduces to the Fermi one for $a=1$. In particular, for $T=0$ and $d=1/a$ the Gentile and Lindsay generalized statistics are identical and therefore all objections concerning the reduction of the nucleon maximum momentum still hold.

In conclusion, generalized statistics, although considered as crude devices allowing a rough calculation of average quantities for aggregates formed by strongly interacting particles, do not seem suitable to describe assemblies like heavy nuclei.

We should like to express our thanks to Professor Nicola Dallaporta and to Doctor Adolfo Kind for their valuable discussions.

1) T. Okayama, Prog. Theor. Phys. **7** (1952), 517.
2) H. S. Green, Phys. Rev. **90** (1953). 270.
3) A. Borsellino, Nuovo Cimento **4** (1947), 52.
4) G. Gentile, Nuovo Cimento **10** (1940), 33; Nuovo Cimento, **4** (1942), 111.
5) J. Bardeen, Phys. Rev. **51** (1937), 799.
6) F. Ferrari and C. Villi, Nuovo Cimento **10** (1953), 94.
7) R. E. Marshak, *Meson Physics*, (McGraw Hill Co, 1952), page 84.
8) G. F. Chew and J. L. Steinberger, Phys. Rev. **78** (1950), 497.
9) R. B. Lindsay, Phil. Mag. **17** (1934), 264.

On the Weak Universal Boson-Fermion Interaction

Shuzo Ogawa,† Hisaichiro Okonogi†
and Sadao Ôneda††

*Institute of Theoretical Physics,
Nagoya University† and Kanazawa University††*

March 18, 1954

The progress of high energy physics has revealed that there exist various sorts of unstable particles and complicated interactions among them. One way of attacking these situations will be to introduce the concept of families of elementary particles and to reduce the individual couplings to a certain kind of interaction between families.[1] In these respects, it will have some significance to search phenomenologically for the interactions which seem to have certain universality and similarity.* Though it is not yet decided which types of interactions are primary, the following interesting facts are known. Before entering into details, we simplify our notations. Write nucleon and Hyperon (Λ^0, $\Lambda^\pm$) by the letter N_0 and N_1 with masses M_0 and M_1. π–meson and θ–meson (summarizing $\tau^\pm$, V_4^0, $\varkappa$ etc.) are denoted by π_0 and π_1 with masses μ_0 and μ_1 and for electron and μ–meson masses we use m_0 and m_1.

(I)   Universal Fermi Interactions.

This standpoint is to assume common interactions of the form $\sum_j f_j (\bar{\psi}_a O_j \psi_b) (\bar{\psi}_c O_j \psi_d)$ among the whole Fermions and to require some selection rules (such as the conservation of heavy charge etc.) for the processes which are not realized in nature. It seems quite natural to explain β–decay, $\mu^\pm$–decay and μ^-–capture by nuclei by these universal interactions.**

(II)   Strong Boson-Fermion Interactions.

It may be possible that besides famous $N_0 N_0 \pi_0$ interactions, $N_0 N_1 \pi_1$ (responsible for the production of N_1 and π_1) and $\bar{N}_1 N_1 \pi_0$ form this category.[2]

* The interactions between charged particles and electro-magnetic field may be regarded to belong to this category.

** If we extend this interaction to comprise Hyperons, we must check, for example, the processes such as $\Lambda^0 \to P + \mu^- (e^-) + \nu$ which compete with observed $\Lambda^0 \to P + \pi^-$. This problem will be discussed in a subsequent letter.

As we have weak interactions such as $\pi_0\pm-$ and Λ–decay, we shall ask in this paper whether there exists the possibility to introduce the other interactions in addition to the possible universal ones (I) and (II) ; that is.

(III)   The weak universal Fermion-Boson interactions.

(A) π^0–meson-lepton interactions.

In setting up this interaction it is instructive to note, first of all, the conspicuous facts that the ratio of the possible π_0–decay life-time $R = \tau(\pi_0\pm \to \mu^\pm + \nu)/\tau(\pi_0\pm \to e^\pm + \nu) \leq 1/1400$ [3].

Now β–decay and $\bar{N}_0 N_0 \pi_0$ inevitably predict $\pi_0\pm \to e^\pm + \bar{\nu}$ (or ν), and the available volume of the phase space favours $\pi_0 \to e$ about 5.5 times more than $\pi_0 \to \mu$ process. From the phenomenological view point, however, there remain two possible types of couplings[4] consistent with experiments even in our usual interactions.

(1) $g \bar{\psi}_a \gamma_5 \gamma_\mu \psi_b \partial_\mu \phi_0 + \text{c.c.}$

(a and b have the same parity)

(2) $g \bar{\psi}_a \gamma_\mu \psi_b \partial_\mu \phi_0 + \text{c.c.}$

(a and b have opposite parity)

where ψ_a, ψ_b and ϕ_0 denote μ–meson (or electron), neutrino (possibly μ_0-meson) and π_0 respectively. Both (1) and (2) give

$$\frac{1}{\tau}(\pi \to \mu) = (g^2/4\pi) m_1^2 \frac{1}{2} \mu_0 \{1 - (m_1/\mu_0)^2 \}^2$$

$$\frac{1}{\tau}(\pi \to e) = (g^2/4\pi) m_0^2 \frac{1}{2} \mu_0 \{1 - (m_0/\mu_0)^2 \}^2.$$

For $\mu_0 = 273\, m_0$, $m_1 = 207\, m_0$ and $\tau(\pi_0 \to \mu) = 2.5 \times 10^{-8}$ sec.,[5] we get $(g^2/4\pi)\mu_0^2 = 3.6 \times 10^{-15}$. These cases give the acceptable ratio $R = 1.3 \times 10^{-4}$. Note that other couplings ("γ_5" or "1") lead to the completely wrong value $R \simeq 5.5$.

(B) π_0–meson-heavy particle interaction.

The life times of Λ–decay suggest that $N_0 N_1 \pi_0$ is weak. Here we consider recently established decay $\Lambda^0 \to P + \pi_0^-$. From (A) we are lead to the universal interaction of the type

(3) $g \psi_\Lambda \gamma_5 \gamma_\mu \psi_P \partial_\mu \phi_0 + \text{c.c.}$

(Λ, $P \cdots$ same parity),

(4) $g \psi_\Lambda \gamma_\mu \psi_P \partial_\mu \phi_0 + \text{c.c.}$

(Λ, $P \cdots$ different parity).

For (3)

$$\frac{1}{\tau} = (g^2/4\pi)(M_1 + M_0)^2 \frac{M_1}{4}$$
$$\times \left\{ \frac{(M_1 + M_0)^2 - \mu_0^2}{M_1^2} \right\}^{1/2} \left\{ \frac{(M_1 - M_0)^2 - \mu_0^2}{M_1^2} \right\}^{3/2}$$

For (4)

$$1/\tau = (g^2/4\pi)\,(M_1 - M_0)^2 \cdot \frac{M_1}{4}$$

$$\times \left\{ \frac{(M_1 + M_0)^2 - \mu_0^2}{M_1^2} \right\}^{3/2} \left\{ \frac{(M_1 - M_0)^2 - \mu_0^2}{M_1^2} \right\}^{1/2}.$$

Taking recent values[6] $\Lambda^0 \to P + \pi^- + Q$ ($Q \sim 40$ Mev) and $\tau = 3 \times 10^{-10}$ sec., we get for (3) $(g^2/4\pi)\,\mu_0^2 = 1.8 \times 10^{-14}$ and for (4) $(g^2/4\pi) \cdot \mu_0^2 = 7.3 \times 10^{-15}$. These values seem to be rather consistent with g in case (A).[7]*

(C) θ-meson-lepton interactions

As regards θ-mesons, their decay schemes have not yet been settled[6] and among their secondaries $\mu^{\pm}$-mesons have sometimes been reported. If we take spin 0[6] for θ, we may extend our universal interactions to include θ. For example, same interactions (replacing π_0 by π_1) as (1) and (2) give for $\theta^{\pm} \to \mu^{\pm} + \nu$ (or μ_0), $\tau\,(\theta \to \mu) = 1.4 \times 10^{-9}$ sec. with $\mu_1 = 970\,m_0$. In this case, it is also remarkable that $\theta^{\pm} \to e^{\pm} + \nu$ is about 2.5×10^{-5} times less rare event. If unique $\theta^0 \to \pi_0^+ + \pi_0^-$ is confirmed, we must forbid the most threatening modes $\theta^0 \to \mu^+ + \mu^-$, $\theta^0 \to \mu^{\pm} + e^{\mp}$ etc. which will have the life times of the order 10^{-9} sec. The identifications of the secondary particles will settle these problems. Note also that other Pais's odd interactions[1] $\bar{N}_0 N_0 \pi_1$ and $\bar{N}_1 N_1 \pi_1$ can also be included in this scheme without serious contradictions.

(D) Consistency with other phenomena.

(α) β-decay. As we introduce direct $\pi_0^{\pm} \to e^{\pm} + \nu$ interaction we must examine β-decay occuring through $(n,\,p)\overset{G}{\to}\pi_0\overset{g}{\to}(e,\,\nu)$. In this scheme free neutron decay life-time τ_β is connected with $\tau\,(\pi_0 \to e)$ via g-interaction ((1) or (2)) by the formula[9]

$$1/\tau_\beta = \left(-\frac{13.8}{\pi} \right) \left(\frac{m_0}{\mu_0} \right)^5 \left(\frac{m_0}{M_0} \right)^2 \left(\frac{G^2}{4\pi} \right)$$

$$\times \,(1/\tau\,(\pi_0 \to e)), \quad G = G_1 + \frac{2M_0}{\mu_0}\,G_1.$$

G_1 is the usual π_0-nucleon PS coupling constant. For $G_1^2/4\pi \sim 1$ and g value in (A), $\tau_\beta \simeq 1.8 \times 10^{10}$ minutes. Thus g interaction will not modify the interpretation of β-decay by the universal Fermi interactions (I). Of course $\bar{N}_0 N_0 \pi_1$ never does affect β-decay by the process $(n,\,p)\overset{g}{\to}\pi_1\overset{g}{\to}(e\nu)$.

(β) μ-capture by nuclei.[7] μ^--capture by nuclei will occur via $(\mu\nu)\overset{g}{\to}\pi_0\overset{G}{\to}(n,\,p)$, and the life time

τ_c is given by,

$$1/\tau_c \simeq 1.2\,G^2\,(a\,Z)^3 \cdot Z \cdot 10^8/\text{sec}.$$

with g value in (A). For $G^2 \simeq 1$, this is consistent with the experiment, and is in the same order as that expected from the universal Fermi interactions. And so, if there exist both processes (I) and (III), one may expect no serious contradictions. Also $\bar{N}_0 N_0 \pi_1$ has no effect on this reaction.

(γ) Λ^0-decay by $\Lambda^0 \overset{G}{\to} P + \theta^- \overset{g}{\to} P + \mu\,(e) + \nu$ etc. The calculated life-times are,

$$\tau\,(\Lambda^0 \to P + \mu^- + \nu) = 3.0 \times 10^{-10}\ \text{sec} \times 10^4$$

and

$$\tau\,(\Lambda^0 \to P + e^- + \nu) = 3.0 \times 10^{-10}\ \text{sec} \times 10^9,$$

respectively, where G and g are the same type couplings as the above. Thus, also in these cases, g interaction will not modify the decay process $\Lambda^0 \to P + \pi^-$.

(E) Speculation on strong and weak universal interactions.

$\bar{N}_0 N_0 \pi_0$ is known to be $PS\,(\gamma_5)$ or $PV\,(\gamma_5 \gamma_\mu)$. If θ^0 (which appear in cosmotoron experiments[6] as $\pi^- + P \to \Lambda^0 + \theta^0$) has a decay mode $\theta^0 \to \pi^+ + \pi^-$, θ^0 must be scalar or vector meson. If θ is scalar, taking the parity of N_1 and N_0 for space inversion as opposite, strong interactions $\bar{N}_0 N_1 \pi_1$ and $\bar{N}_1 N_1 \pi_0$ will also be able to take the form "γ_5" or "$\gamma_5 \gamma_\mu$". On the other hand, weak $\bar{N}_0 N_0 \pi_1$, $\bar{N}_0 N_1 \pi_0$ and $\bar{N}_1 N_1 \pi_1$ may all have the same type "γ_μ" (not "1", "γ_5", "$\gamma_5 \gamma_\mu$") which might be extended to comprise also lepton $-\pi_0$ and $-\pi_1$ interactions if we assign appropriate parity for leptons.

(F) Possible reasons for the necessity of the weak Universal Boson-Fermion Interaction.

There may remain the possibility that only universal interactions of the form (I) and (II) between families of elementary particles are enough to realize all observed processes. That is, for example, $N_1 \to N_0' + \pi_0$ may occur through $N_1 \overset{f}{\to} \bar{N}_0 + N_0 + N_0'$ and as is well-known, $\pi_0^{\pm} \to \mu^{\pm} + \nu$ may be explained by $\pi_0 \to N_0 + \bar{N}_0' \overset{f}{\to} \mu + \nu$. To our regret, this proposition, accompanied by serious divergence difficulty, cannot be traced correctly. In the latter case, however, we must also ask why $\pi_0^{\pm} \to \mu^{\pm} + \nu$ predominates over $\pi_0^{\pm} \to e^{\pm} + \nu$. In this case the divergences will cancel in the ratio R and we may have some conclusions. Naming Fermi couplings by the Dirac Matrices O in $(\bar{\psi}_N O \psi_N)$, the lowest

---

* In (3) and (4) both Λ^0 and nucleon are assumed to be normal particles. If Λ^0 is antiparticle (hole) and nucleon is normal the results of (3) and (4) must be interchanged.

order processes are forbidden for S, T and V.* For PS the ratio $R \simeq 5.5$ and only PV coupling gives the right order 1.3×10^{-4}. This is due to the fact that PV coupling effectively leads to the g coupling of the form (1) or (2). Thus the explanation of R by both g coupling (1) or (2) and by Universal PV Fermi coupling must be regarded to be on the same basis. If PV Fermi coupling does not play a predominant role the necessity of our universal Boson-Fermion interactions will be increased.** In these points, the determination of the types of the universal Fermi Interactions is enlightening. We wish to express our hearty thanks to the group of Nagoya, and Kanazawa University for their helpful discussions.

1) A. Pais, Phys. Rev. **86** (1952), 633. See also S. Oneda Prog. Theor. Phys. **9** (1953), 327, especially 4.

2) As the incorporation of this idea, see, for example, A. Pais, Prog. Theor. Phys. **10** (1953), 457. Also for the Universal Fermi Interactions, see O. Hara, T. Marumori, Y. Ohnuki and H. Shimodaira, ibid. **10** (1953), 470.

3) H. Friedman and J. Rainwater, Phys. Rev. **81** (1951), 644. F. M. Smith, Phys. Rev. **81** (1951), 897.

4) The particular behaviour of these couplings was first noted by S. Sakata. Proc. Phys-Math. Soc. **23** (1941), 291. See also B. D. Espagmat, Compt. Rend. **228** (1948), 749.

5) R. P. Durbin, H. H. Loar and W. Haveus, Jr. Phys. Rev. **88** (1952), 179.

6) Reports of the international conference held in Kyoto Sep. 1953.

7) For PV coupling, see S. Ogawa ; " Soryushiron-kenkyu " (mimeographical circular in Japanese) **5** (1953), 895 ; 1212.

8) Spin 0 for θ has the advantage that γ-decay, $\pi_1 \to \pi_0 + \gamma$, is absolutely forbidden.

9) E. R. Caianiello, Phys. Rev. **81** (1951), 625.

10) I. Fujita and M. Yamada, Prog. Theor. Phys. **10** (1953), 518.

* When photon is included these also contribute to this process, but they will also favour $\pi_0 \to e$ more than $\pi_0 \to \mu$.

** When PV coupling does not appear in Fermi interactions or when PS coupling exists with appreciable magnitude (in this case we must take the viewpoint that π_0-decay through nucleon loop will have small effect, though we cannot evaluate it correctly at present), we must have recourse to direct g coupling for $\pi_0 \pm$-decay. At present, however, the cases $S - T$ with $PV (g_{PV} < (g_S$ or $g_T)/(10-30))$ and $PS (g_{PS} < 10^{-2} g_{PV})$ seems not to be rejected, and there remains the possibility that this PV gives the correct R.[10] We are indebted to M. Umezawa for his useful discussions.

Kinematical Studies of Pion-Nucleon Interactions

Satio Hayakawa,* Masaaki Kawaguchi* and Shigeo Minami**

Institute for Fundamental Physics, Kyoto University* *and Department of Physics, Osaka University***

March 29, 1954

It has been pointed out by one of us (S. M.)[1] that the angular distribution as well as its absolute magnitude for pion nucleon scattering are invariant under the simultaneous interchange of every pair of phase shifts belonging to the same total angular momentum. In view of the widely spread belief that the pion-nucleon interaction is particularly strong in $f_{3/2}$ state, and also that the strength of the spin-orbit coupling for a nucleon in nuclei is estimated with the doublet splitting observed in the scattering of nucleons by even-even nuclei, we think it worth while to make a closer investigation on the reaction between spin 0 and spin 1/2 particles.

First of all, we would like to mention that the degeneracy pointed out in I is implied in the Racah coefficient appearing in the formula of the angular distribution.[2] There are two symmetry properties of the Z-coefficient :

$$Z\left(J - \frac{1}{2}, J, J' - \frac{1}{2}, J' ; \frac{1}{2}, L\right)$$
$$= Z\left(J + \frac{1}{2}, J, J' + \frac{1}{2}, J' ; \frac{1}{2}, L\right), \quad (1a)$$
$$Z\left(J + \frac{1}{2}, J, J' - \frac{1}{2}, J' ; \frac{1}{2}, L\right)$$
$$= -Z\left(J - \frac{1}{2}, J, J' + \frac{1}{2}, J' ; \frac{1}{2}, L\right). \quad (1b)$$

Here we use the notations in reference (2). Since there appear in the differential scattering cross section under consideration only the squares of Z's such that given in (1a) and (1b), these relations give us the general proof equivalent to, but different in appearance from that in I. We emphasize that the degeneracy concerned exists in any reaction if it takes place only in channel spin 1/2. The above relations are useful to find out the inter-relations among partial waves, when there is the channel spin 1/2 in part.

To remove the above degeneracy in the cross section, we suggest to observe the polarization of

recoil nucleons in the pion-nucleon scattering, or of scattered nucleons in the $n-\alpha$ scattering. The observation of the polarization of the recoil nucleon has been suggested by Fermi[3] so as to distinguish Fermi's phase shifts from Yang's. This is useful also in our case, as the degeneracy arises from the sum over possible spin states. The polarization can be derived, in reference to the general formula for the unpolarized initial state, as

$$P = \frac{AB^* + BA^*}{AA^* + BB^*}\, \boldsymbol{n}, \qquad (2)$$

where $\boldsymbol{n}$ is the unit vector normal to the plane in which the scattering occurs, as defined by Lepore.[4] And

$$AB^* + BA^* = \sin\theta \sum_{l,\,l'}(l+1)\,(l'+1)\,i$$

$$\times \{(R_l^+ R_{l'}^{+*} - R_{l+1}^- R_{l'+1}^{-*})\,(P_l P_{l'+1} - P_{l+1} P_{l'})$$

$$+ (R_l^+ R_{l'+1}^{-*} - R_{l+1}^- R_{l'}^{+*})\,(P_l P_{l'} - P_{l+1} P_{l'+1})\},$$

$$(3a)$$

$$AA^* + BB^* = \sum_{l,\,l'}(l+1)\,(l'+1)$$

$$\times \{(R_l^+ R_{l'}^{+*} + R_{l+1}^- R_{l'+1}^{-*})$$

$$\times [P_l P_{l'} + P_{l+1} P_{l'+1} - \cos\theta\,(P_l P_{l'+1} + P_{l+1} P_{l'})]$$

$$+ (R_l^+ R_{l'+1}^{-*} + R_{l+1}^- R_{l'}^{+*})$$

$$\times [P_l P_{l'+1} + P_{l+1} P_{l'} - \cos\theta\,(P_l P_{l'} + P_{l+1} P_{l'+1})]\}. \qquad (3b)$$

P_l is the Legendre function and R_l^+ and R_l^- are the R-matrices with regard to the states of $J = l \pm 1/2$, corresponding to e_J and f_{J-1} in I, respectively. From this one can see that the sense of the polarization is just upset by the interchange of phase shifts of $l = J + 1/2$ and $l = J - 1/2$, as can be expected.

For convenience of analyzing experiments we give the polarization in terms of phase shifts up to $J = 3/2$.

$$AB^* + BA^* = 2\sin\theta\,[\sin\delta_0^+ \sin\delta_1^- \sin(\delta_1^- - \delta_0^+)$$

$$+ \sin\delta_1^+ \sin\delta_0^+ \sin(\delta_0^+ - \delta_1^+)$$

$$+ \sin\delta_1^- \sin\delta_2^- \sin(\delta_2^- - \delta_1^-)$$

$$+ 3\cos\theta\,\{\sin\delta_0^+ \sin\delta_2^- \sin(\delta_2^- - \delta_0^+)$$

$$+ \sin\delta_1^+ \sin\delta_1^- \sin(\delta_1^- - \delta_1^+)\}$$

$$+ (1 - 9\cos^2\theta)\sin\delta_2^- \sin\delta_1^+ \sin(\delta_1^+ - \delta_2^-)],$$

$$AA^* + BB^* = \sin^2\delta_0^+ + \sin^2\delta_1^-$$

$$+ 2\cos\theta \sin\delta_0^+ \sin\delta_1^- \cos(\delta_0^+ - \delta_1^-)$$

$$+ 4\cos\theta\,\{\sin\delta_0^+ \sin\delta_1^+ \cos(\delta_0^+ - \delta_1^+)$$

$$+ \sin\delta_1^- \sin\delta_2^- \cos(\delta_1^- - \delta_2^-)\}$$

$$+ 2\,(3\cos^2\theta - 1)\,\{\sin\delta_0^+ \sin\delta_2^- \cos(\delta_0^+ - \delta_2^-)$$

$$+ \sin\delta_1^- \sin\delta_1^+ \cos(\delta_1^- - \delta_1^+)\}$$

$$+ (1 + 3\cos^2\theta)\,(\sin^2\delta_1^+ + \sin^2\delta_2^-)$$

$$+ 2\,(9\cos^3\theta - 5\cos\theta)$$

$$\times \sin\delta_1^+ \sin\delta_2^- \cos(\delta_1^+ - \delta_2^-). \qquad (4)$$

Full accounts of the kinematical analysis of pion-nucleon reactions, including radiative processes, will be published soon.

1) S. Minami, Prog. Theor. Phys. **11** (1954), 213, referred to as I.
2) J. M. Blatt and L. C. Biedenharn, Rev. Mod. Phys. **24** (1952), 258.
3) E. Fermi, Phys. Rev. **91** (1953), 947.
4) J. V. Lepore, Phys. Rev. **79** (1950), 137.
   We take the sense of $\boldsymbol{n}$ opposite to Lepore's, so that our sign of polarization is the same as Fermi's.

The Effect of the Mass Difference between Charged and Neutral Pions on the Nuclear Force

Atsushi Sugie

Department of Physics, University of Tokyo

March 31, 1954

It has long been well known that there is a small difference between the scattering length in the ${}^1S$ state for the $n-p$ system a_{np} and that for the $p-p$ system a_{pp} even after the effect of the Coulomb force is eliminated. The effect of the Coulomb force differs a little for each assumed shape of the nuclear force but, under reasonable well shapes, a_{np} does not become equal to a_{pp}.

It was shown by Schwinger[1] that the electromagnetic interaction due to the anomalous magnetic moment of proton and neutron can account for this difference if the Yukawa well is assumed for the

Table 1. $'b = \hbar^2/Mc^2$

Core radius (10^{-13} cm)	0.0	0.3	0.6
b/a_{np}	1.36	1.28	1.24
b/a_{pp}	1.39	1.65	1.78

(The potential outside the core is the Hulthén potential)

nuclear force. But recently, Salpeter[2] demonstrated that when a high repulsive core (say $1/2$ Mc^2) is introduced in the potential for the small nucleon distance, a slight difference still remains between a_{np} and a_{pp}. His result is quoted in the table 1.

We know that the nuclear force derived from the charge symmetric meson theory is the same for both n–p and p–p system. But in reality, the mass of the charged pion is a little larger than that of the neutral, and we can take into account this mass difference semi-empirically, thus making the nuclear force not charge independent. Under the ps–pv theory or the ps–ps theory with the completely damped pair term,[3] the charge dependent term thus obtained is, including the fourth order term,

$$\tau_1{}^z \tau_2{}^z \Delta\mu \left\{ \frac{G^2}{4\pi}\left(\frac{\mu}{2M}\right)^2\left(\frac{2}{x}-1\right)e^{-x}\right.$$

$$+\left(\frac{G^2}{4\pi}\right)^2\left(\frac{\mu}{2M}\right)^4 \frac{8}{\pi}\left[\frac{K_0(2x)}{x}\right.$$

$$\left.\left.+\left(1+\frac{11}{4x^2}\right)K_1(2x)\right]\right\}$$

where $x = \mu r$ and

$$\Delta\mu/\mu = [\mu(\text{charged}) - \mu(\text{neutral})]/\mu = 1/30 .[4]$$

In the usual ps–pv interaction Hamiltonian,

$$(g/\mu)\,\bar\psi\,(\sigma\nabla)\,\tau_\alpha\,\psi\,\phi_\alpha$$

the μ under g is introduced only to make g dimensionless and has no physical significance. We replace g/μ with $G/2M$ and then the above expression follows. There is an argument concerning the treatment of the velocity dependent term but its contribution is negligible for the ^1S state. Our result corresponds to the older treatment.[5] The value of $G^2/4\pi$ and the core radius were determined from a_{np} and the singlet effective range,[5]

$$G^2/4\pi = 16, \quad r_{\text{core}} = 0.54 \times 10^{-13} \text{ cm (hard core)}.$$

Using these values and the wave function at zero energy

$$u = 1 + x/a - Ae^{-\beta x}, \quad \text{for } x \geq x_{\text{core}},$$

$$u = 0 \qquad \qquad \text{for } x \geq x_{\text{core}},$$

we have

$$b/a_{pp} - b/a_{np} = 0.40 ,$$

which is quite reasonable compared with the table 1.

The mass difference, as well as the anomalous

magnetic moment, is the consequence of the interaction of the pion and the nucleon field with the photon field. This interaction will also make the coupling constant G different for the charged and the neutral pion. But the difference is not yet determined experimentally nor can be predicted from the present theory. The charge dependent part of the nuclear force is more sensitive to the difference of G than that of μ. This is due to the fact that if the mass becomes smaller, the force range grows longer and the strength weaker. These two effects partially cancel each other and the net result is not large. On the contrary, the change of the coupling constant alters merely the strength of the force and the contribution from the fourth order term is quite large. The effect of the mass difference would be canceled completely if the coupling constant for the neutral pion should be smaller by about 1%.

········ pion line
⎓⎓⎓ photon line
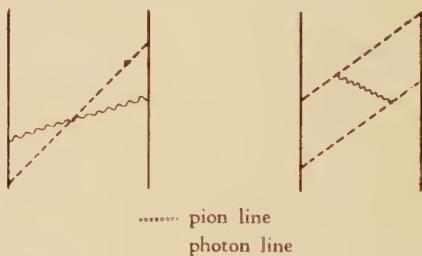
Furthermore, there are charge dependent potentials which can not be attributed to the change of μ and G, such as the ones illustrated by the diagrams, but these effects will be very small.

While these contributions are not yet calculated, we can surely conclude that there is nothing unexpected if the specific nuclear forces for the n–p and the p–p system have a slight difference.

The author should like to express his sincere thanks to Dr. K. Sawada and Dr. K. Nishijima for their valuable suggestions and to Dr. S. Ohkubo for his stimulating discussions.

1) J. Schwinger, Phys. Rev. **78** (1950), 135.
2) E. E. Salpeter, Phys. Rev. **91** (1953), 994.
3) K. A. Brueckner and K. M. Watson, Phys. Rev. **92** (1953), 1023.
4) W. Chinowsky and J. Steinberger, Phys. Rev. **93** (1954), 586.
5) M. Taketani, S. Machida and S. O-numa, Prog. Theor. Phys. **7** (1952), 45.

On the Neutron He⁴ Scattering

Taro Tamura

*Department of Physics, Tokyo University
of Education*

March 31, 1954

Recently, it has been reported many very accurate experimental results on the total and the differential cross sections for the neutron-He⁴ scattering,[1]−[4] as well as the proton-He⁴ sacttering,[5]−[8] and the phase-shifts of various partial waves were obtained.[2] [4] [9]

To these experimental data, several theoretical approaches were performed, either by taking the shell,[10] [11] or the nuclear reaction theory,[9] [12] or the single particle potential assumption.[13] [14]

The neutron-He⁴ scattering is an important problem, as He⁵ is the lightest nucleus in which marked evidence of the spin-orbit interaction appears, and effects of many body forces, or the deviations of the exchange properties of the two body interactions from those which are known from two body problems may appear.

This is a preliminary report for the theoretical approach in which nuclear forces are assumed explicitly, and the anti-symmetrization among similar particles are taken into account.

We assume here the two body central forces

$$V_{ij}^{(c)} = V_c \, (w + m P_\tau + h P_\tau + b P_\sigma) \exp[-\nu r_{ij}^2] \, ,$$

and the two body spin-orbit interactions[10] [11]

$$V_{ij}^{(s)} = V_s \, (\boldsymbol{\sigma}_i + \boldsymbol{\sigma}_j) \cdot \boldsymbol{L}_{ij} \exp[-\chi r_{ij}^2] \, .$$

Then using the method of the "resonating group structures" of Wheeler,[15] the basic integro-differential equation is obtained as follows:

$$[d^2/dr^2 + k^2 - l(l+1)/r^2] f(r)$$
$$= (4w - 2b - m - 2h) U_c [16a/16a + 3\nu]^{3/2}$$
$$\times \exp[-16a\nu r^2/16a + 3\nu] f(r)$$
$$- (k^2 + 2\mu |E_\alpha|/\hbar^2) [3a/\pi]^{3/2}$$
$$\times \int r\rho \exp[-136a(r^2 + \rho^2)/75]$$
$$\times Q_l (128ar\rho/75) f(\rho) d\rho$$
$$- (4/5) [3a/\pi]^{3/2} \int r\rho [\{188a/5$$
$$- (95/4) (64a/75)^2 (r^2 + \rho^2)\} Q_l (128ar\rho/75)$$
$$- (245/8) (64a/75)^2 r\rho \sum_{\bar{l}} (10\bar{l}0|\bar{l}0)^2$$
$$\times Q_{\bar{l}} (128ar\rho/75)]$$
$$\times \exp[-136a(r^2 + \rho^2)/75] f(\rho) d\rho$$

$$- U_c \Big[(3w + 3m) [6a^2/\pi (2a+\nu)]^{3/2} \int r\rho$$
$$\times \exp[-136a(r^2 + \rho^2)/75] Q_l (128ar\rho/75) f(\rho) d\rho$$
$$+ (3w + 3m) [18a^2/\pi (6a+\nu)]^{3/2} \int r\rho$$
$$\times \exp[-8a\{134a + 7\nu) r^2$$
$$+ (34a + 27\nu) \rho^2\}/25 (6a+\nu)]$$
$$\times Q_l (128a(2a+\nu) r\rho/25 (6a+\nu)) f(\rho) d\rho$$
$$+ (3w + 3m) [18a^2/\pi (6a+\nu)]^{3/2} \int r\rho$$
$$\times \exp[-8a\{(34a + 27\mu) r^2$$
$$+ (34a + 7\nu) \rho^2\}/25 (6a+\nu)]$$
$$\times Q_l (128a(2a+\nu) r\rho/25 (6a+\nu) f(\rho) d\rho$$
$$- (4m + 2h - w + 2b) [3a/\pi]^{3/2} \int r\rho$$
$$\times \exp[-8 (17a + 6\nu) (r^2 + \rho^2)/75]$$
$$\times Q_l (32 (4a - 3\nu) r\rho/45) f(\rho) d\rho \Big]$$
$$- U_{s(-)}^{l+j+1/2} W (1/2 \, 1/2 \, l \, l; 1 \, j) 5 [16a/16a + 3\chi]^{3/2}$$
$$\exp[-16a\chi r^2/16a + 3\chi] f(r)$$
$$- U_{s(-)}^{l+j+1/2} W (1/2 \, 1/2 \, l \, l; 1 \, j)$$
$$\times \sum_{\bar{l}} (2\bar{l} + 1) (10\bar{l}0|l0)^2 W (11 \, \bar{l} \, l; 1 \, l)$$
$$\times 64\pi^{5/2} a^{-3/2} \int r^2 \rho^2 \Big[(9/5 \sqrt{3}) (6a/6a + \chi)^{5/2}$$
$$\times \{\exp[-8a\{(34a + 7\chi) r^2$$
$$+ (34a + 27\chi) \rho^2\}/25 (6a+\chi)]$$
$$+ \exp[-8a\{(34a + 27\chi) r^2$$
$$+ (34a + 7\chi) \rho^2\}/25 (6a+\chi)]$$
$$\times Q_l (128 (2a + \chi) ar\rho/25 (6a+\chi))$$
$$+ (3\sqrt{3}/5) \exp[-8 (17a + 6\chi) (r^2 + \rho^2)/75]$$
$$\times Q_{\bar{l}} (32 (4a - 3\chi) r\rho/75) \Big] f(\rho) d\rho \, . \qquad (1)$$

Here μ is the reduced mass, $U_c = 2\mu V_c/\hbar^2$, $U_s = 2\mu V_s/\hbar^2$, $Q_l(A) = 4\pi i^l j_l(iA)$ and j_l is the spherical Bessel function, and we have assumed the space part of the He⁴ nucleus as $N_\alpha^{-1/2} \exp(-a/2 \cdot \sum_{i<j=1}^{4} r_{ij}^2)$, $N_\alpha^{-1/2}$ being the normalization constant.

Experimentally it is known that the $s_{1/2}$-phase shift is negative and its energy dependence is approximately $-ka$, where k is the wave number in c.m.s., and $a = 2.9 \times 10^{-13}$cm. As there are no bound s states in He⁵, this will mean the existence of a rather strong repulsive potential for the s-wave scattering. On the other hand $p_{1/2}$- and $p_{3/2}$-phase shifts are both positive, showing that the effective potentials for the p-states are attractive, although the

spin-orbit interaction makes the potentials shallower and deeper in these two cases, respectively. The difference between the s-phase shift, and the p-phase shifts as a whole, should be explained by central forces only.

It is interesting to note that the main terms of the kernels in (1) are all proportional to Q_l, which is positive for $l=0$, while is negative for $l=1$. Therefore it would be possible, in our formalism, to explain the different behaviour in the s- and the p-phase shifts, by taking appropriate values for nuclear parameters, contrary to the cases of one body potential assumptions,[13] [14] in which no correlation of these two kinds of phase shifts were discussed

Thus it might be expected also, that our formalism will give somewhat different values for nuclear parameters, from those obtained in the one body potential assumptions.

Numerical calculations are being done, and the results will be reported in this journal presently.

1) S. Bashkin, F. P. Mooring and B. Petree, Phys. Rev. **82** (1951), 378.

2) P. Huber and E. Baldinger, Helv. Phys. Acta. **25** (1952), 435.

3) J. H. Coon, E. R. Graves and H. H. Barshall, Phys. Rev. **88** (1952), 562.

4) J. D. Seagrave, Phys. Rev. **92** (1953), 1222.

5) G. Freier, E. Lampi, W. Sleater and J. H. Williams, Phys. Rev. **75** (1949), 1345.

6) M. Heusinkveld and G. Freier, Phys. Rev. **85** (1952), 80.

7) W. E. Kreger, R. O. Kerman and W. K. Jentschke, Phys. Rev. **86** (1952), 593.

8) T. M. Putman, Phys. Rev. **87** (1952), 932.

9) D. C. Dodder and J. I. Gammel, Phys. Rev. **88** (1952), 520.

10) J. Hughes and K. J. Le Couteur, Proc. Phys. Soc. **63** (1950), 1219.

11) C. H. Blanchard and R. Avery, Phys. Rev. **81** (1951), 35.

12) R. K. Adair, Phys. Rev **86** (1952), 155.

13) B. Bürgel, referred to in reference 2).

14) S. Sack, L. C. Biedenharn and G. Breit, Phys. Rev. **93** (1954), 321.

15) J. A. Wheeler, Phys. Rev. **52** (1937), 1107.

On the Theory of Condensing Systems*

Kazuyoshi Ikeda

Physics Department, Faculty of Science, Kyusyu University

March 29, 1954

We give a method by which one treats the Ursell expansion[1] of the configuration integral of the system of N interacting particles

$$Q_N = \sum_{m_l \geq 0} \left(\sum_{l=1}^{l=N} l m_l = N\right) \prod_{l=1}^{l=N} \frac{(Vb_l)^{m_l}}{m_l!} . \quad (1)$$

Here b_l are the cluster integrals defined by

$$b_l = \frac{1}{l! \, V} \int \sum_{1}^{\text{cluster}} \prod_{l > i > j > 1} f_{ij} d\{l\}, \quad (2)$$

where

$$f_{ij} = f(r_{ij}) = \exp\left\{-\frac{1}{kT}\phi(r_{ij})\right\} - 1, \quad (3)$$

$\phi(r_{ij})$ being a pair potential.

Our method of treatment is similar to Mayer's method[2] of taking the maximum term, but the discussion of its difference from that of Mayer in essential points will be given later.

§ 1. Method of treatment

We consider the following limit: Put $V = Nv$, and make $N \to \infty$, v being fixed.

As $N \to \infty$, we have quantities X and x such that

$$\left.\begin{array}{l} \lim_{N \to \infty} X/N = \text{definite and finite and positive,} \\ \lim_{N \to \infty} x/N = 0. \end{array}\right\} \quad (4)$$

We call such X "large", and such x "small", and simply write $X \sim N$, $x \ll N$. Then we mean by $a \sim 1$ that $0 < \lim_{N \to \infty} a < \infty$, and by $M \gg 1$ that $\lim_{N \to \infty} M = \infty$. It follows that all such X are ≥ 1, but some x are ~ 1 and other x are ≥ 1.

Now in (1), on making $N \to \infty$, we may divide all l into "small" l and "large" l, which we call l and L respectively. (i.e. $l \ll N$, $L \sim N$). Then the integrals b_l and b_L are called "small" cluster integrals and "large" cluster integrals respectively, and the corresponding clusters of molecules are called "small" clusters and "large" clusters. We see that the number m_l of clusters of l molecules

satisfies $m_l \gg 1$ or $m_l \sim 1$ or $m_l = 0$ while the number m_L of clusters of L molecules satisfies $m_L \sim 1$ or $m_L = 0$.

Let N_1 be the total number of those molecules which compose "small" clusters, and N_2 be the total number of those molecules which compose "large" clusters. Then we have

$$\sum_{l=1}^{l \ll N} lm_l = N_1, \quad \sum_{L \sim N}^{L=N} Lm_L = N_2, \quad N_1 + N_2 = N. \quad (5)$$

Hence we may write (1) as

$$\Omega_N = \sum_{N_2, m_l} (\sum_{l=1}^{l \ll N} lm_l + N_2 = N) \prod_{l=1}^{l \ll N} \frac{(Nvb_l)m_l}{m_l!}$$

$$\times \sum_{m_L} (\sum_{L \sim N}^{L=N_2} Lm_L = N_2) \prod_{L \sim N}^{L=N_2} \frac{(Nvb_L)m_L}{m_L!}. \quad (6)$$

We should treat the first terms $\prod_{l=1}^{l \ll N} \cdots$ and the last terms $\prod_{L \sim N}^{L=N_2} \cdots$ in different ways, because of the difference of their properties. Here we assume that $b_l > 0$, $b_L > 0$ for all l, L, and assume for the present that b_l, b_L are independent of volume. Then we have

$$\ln \sum_{m_L} (\sum_{L \sim N}^{L=N_2} Lm_L = N_2) \prod_{L \sim N}^{L=N_2} \frac{(Nvb_L)m_L}{m_L!}$$

$$\stackrel{\ldots}{=} \ln \prod_{L \sim N}^{L=N_2} \frac{(Nvb_L)m_L}{m_L!}$$

$$\stackrel{\ldots}{=} \ln(NvbN_2) \stackrel{\ldots}{=} \ln bN_2 \stackrel{\ldots}{=} N_2 \ln b_0, \quad (7)$$

where by the equality symbol $\stackrel{\ldots}{=}$ we mean that on both sides of the equation all "small" quantities have been neglected. Here we used the fact that $\lim b_L{}^{1/L} = b_0$ and hence $b_L = f(L)b_0{}^L$ where $\ln f(L) \ll L \sim N$ that is $\ln f(L)$ is "small", and that $m_L \sim 1$, $\sum_{L \sim N}^{L=N_2} 1 \sim 1$, and that the logarithm of the total number of the terms in $\sum_{m_L}$ is "small". Next we have

$$\ln \prod_{l=1}^{l \ll N} \frac{(Nvb_l)m_l}{m_l!} \stackrel{\ldots}{=} \sum_{l=1}^{l \ll N} m_l(\ln Nvb_l - \ln m_l + 1), \quad (8)$$

because we can apply Stirling's approximation formula to all $m_l \gg 1$, and the existence of those terms for which $l \ll N$ and $m_l \sim 1$ is proved to be negligible.

Thus we have

$$\Omega_N = \sum_{N_2, m_l} (\sum_{l=1}^{l \ll N} lm_l + N_2 = N) T(N_2, m_l), \quad (9)$$

where

$$\ln T(N_2, m_l) \stackrel{\ldots}{=} \sum_{l=1}^{l \ll N} m_l(\ln Nvb_l - \ln m_l + 1) + \ln bN_2$$

$$\stackrel{\ldots}{=} \sum_{l=1}^{l \ll N} m_l(\ln Nvb_l - \ln m_l + 1) + N_2 \ln b_0. \quad (10)$$

Since the logarithm of the total number of the terms in $\sum_{N_2, m_l}$ of (6) is "small", we have

$$\ln \Omega_N \stackrel{\ldots}{=} \text{Max} \ln T(N_2, m_l) = \ln T(N_2^*, m_l^*). \quad (11)$$

Here the maximum is taken under the condition

$$\sum_{l=1}^{l \ll N} lm_l + N_2 = N, \quad (12)$$

and so, using Lagrange's multiplier $\ln z$, we obtain

$$z = b_0{}^{-1} (\equiv z_s) \quad (13)$$

from $\partial/\partial N_2 \{ \ln T(N_2, m_l) + (\sum_{l=1}^{l \ll N} lm_l + N_2) \ln z \} = 0$, and obtain

$$m_l^* = Nvb_l z^l = Nvb_l z_s{}^l \quad (l = 1, 2 \cdots \ll l) \quad (14)$$

from $\partial/\partial m_l \{ \quad \} = 0$ $(l = 1, 2, \cdots \ll l)$. Hence we know that m_l^* is "large" for every $l \sim 1$. The values of N_1 and N_2 for the maximum term are

$$N_1^* = \sum_{l=1}^{l \ll N} lm_l^* = N \sum_{l=1}^{l \ll N} lvb_l z_s{}^l \quad (15)$$

and

$$N_2^* = N - N \sum_{l=1}^{l \ll N} lvb_l z_s{}^l \quad (16)$$

respectively.

Then, in order that the solution just now obtained may have a physical meaning, it is necessary that $\sum_{l=1}^{l \ll N} lvb_l z_s{}^l$ should be definite and finite and $\sum_{l=1}^{l \ll N} lvb_l z_s{}^l \leq 1$. In order that this may be so, firstly, the series $\sum_{l=1}^{\infty} lb_l z_s{}^l [= \sum_{l=1}^{\infty} lf(l)$ where $b_l = f(l)b_0{}^l]$ should be convergent (notice that some $l \ll N$ are $\gg 1$), unless v is infinitesimal. (In this case, obviously, $\sum_{l=1}^{l \ll N} lb_l z_s{}^l = \sum_{l=1}^{\infty} lb_l z_s{}^l$, and $\sum_{l=1}^{l \ll N} b_l z_s{}^l [= \sum_{l=1}^{\infty} f(l)]$ is convergent and is equal to $\sum_{l=1}^{\infty} b_l z_s{}^l$). Secondly, the following should hold:

$$\frac{1}{v} \geq \sum_{l=1}^{\infty} lb_l z_s{}^l \left(\equiv \frac{1}{v_s} \right). \quad (17)$$

Under the above conditions we obtain

$$\ln \Omega_N \stackrel{\ldots}{=} N \sum_{l=1}^{l \ll N} vb_l z_s{}^l - N_1^* \ln z_s + \ln bN_2^* \quad (18)$$

with (15) and (16). Here we can replace the upper limit of the sum $l \ll N$ by $l \to \infty$, and (18) becomes, using $z_s \equiv b_0{}^{-1}$,

$$\ln \Omega_N \stackrel{\ldots}{=} N \sum_{l=1}^{\infty} vb_l z_s{}^l - N \ln z_s. \quad (19)$$

Hence the pressure of the system is

$$p = kT \left(\frac{\partial \ln \Omega_N}{\partial V} \right) = kT \sum_{l=1}^{\infty} b_l z_s{}^l (\equiv p_s), \quad (20)$$

which is independent of v. And the Gibbs free energy is

$$G = F + pV \quad \text{(where } F = -kT \ln Q_N, \quad Q_N = \lambda^{-3N} Q_N\text{)}$$

$$\rightleftharpoons NkT \ln \lambda^3 z_s, \quad \text{where } \lambda \equiv h/\sqrt{2\pi mkT}, \quad (21)$$

and so z_s is the fugacity of the system, which is independent of v. This corresponds to the condensation range.

On the other hand, if v is given so that

$$\frac{1}{v} < \sum_{l=1}^{\infty} l b_l z_s{}^l \left(\equiv \frac{1}{v_s} \right), \tag{22}$$

then we have $N_2{}^* < 0$ in (16) and so we should have $N_2{}^* = 0$ for the greatest term which has a physical meaning. In this case, we have

$$\ln \Omega_N \rightleftharpoons \text{Max} \ln T(0, m_l) = \ln T(0, m_l) \tag{23}$$

with the condition

$$\sum_{l=1}^{l \ll N} l m_l = N. \tag{24}$$

Using Lagrange's multiplier $\ln z$, we obtain

$$m_l{}^* = N v b_l z^l \quad (l = 1, 2, \cdots \ll N), \tag{25}$$

where z is determined by

$$N \sum_{l=1}^{l \ll N} l v b_l z^l = N. \tag{26}$$

Hence we have

$$\frac{1}{v} = \sum_{l=1}^{l \ll N} l b_l z^l = \sum_{l=1}^{\infty} l b_l z^l : \text{ convergent} \tag{26}$$

unless $v \to 0$, and we have

$$\sum_{l=1}^{l \ll N} b_l z^l \quad \sum_{l=1}^{\infty} b_l z^l : \text{ convergent.}$$

Thus, from (23) and (25), we obtain

$$\ln \Omega_N \rightleftharpoons N \sum_{l=1}^{\infty} v b_l z^l - N \ln z \tag{27}$$

and

$$p = kT \sum_{l=1}^{\infty} b_l z^l \tag{28}$$

and

$$G \rightleftharpoons NkT \ln \lambda^3 z, \tag{29}$$

which shows that z is the fugacity of the system. This corresponds to the gaseous phase.

Thus, v_s [(17), (22)] and z_s [(13)] are the specific volume and the fugacity at the starting-point of condensation, respectively. We see from $\lim_{l=\infty} b_l{}^{1/l} = b_0$ that $z_s (\equiv b_0{}^{-1})$ is the first singularity on the real positive axis of z, of the power series p [(28)] and $1/v$ [(26')]

§ 2. Discussions and remarks

Mayer derived, by his maximum-term method,

$$N = \sum_{l=1}^{l=N} l V b_l z^l \quad \text{or} \quad \frac{1}{v} = \sum_{l=1}^{l=N} l b_l z^l \tag{30}$$

for the condensation range (i.e. the horizontal part of an isotherm) as well as for the gaseous phase. We can show that Mayer's equation (30) is correct only for the gaseous phase, but not for the condensation range, as we see from the fact that the equation $N = V b_1 z + 2 V b_2 z^2 + \cdots + (N-1) V b_{N-1} z^{N-1} + N V b_N z^N$, where $z = z_s e^\varepsilon$ and $\varepsilon (>0)$ is very small, is inconsistent with the thermodynamic rule of extensive variables. For, in Mayer's treatment, the last terms in (6) were not distinguished from the first terms when they were treated, and so the greatest term was not correctly taken in the condensation range, where the last terms become important.

Equation (30) is also obtained by the saddle-point (steepest-descent) method, but this method is proved to be valid only for the gaseous phase, but not for the condensation range, which is consistent with the above statement.

On the other hand, by our method, in which the above-mentioned faults may be rectified, we have obtained

$$N = \sum_{l=1}^{\infty} l V b_l z^l \quad \text{or} \quad \frac{1}{v} = \sum_{l=1}^{\infty} l b_l z^l \tag{31}$$

for the gaseous phase (see (26)), and

$$N = \sum_{l=1}^{\infty} l V b_l z_s{}^l + N_2{}^* \quad \text{or} \quad \frac{1}{v} = \sum_{l=1}^{\infty} l b_l z_s{}^l + \frac{1}{v_2} \tag{32}$$

for the condensation range (see (16)), and in particular have derived the last term $N_2{}^*$, which comes directly from the last term $\ln b N_2{}^*$ in (18) or $\ln b N_2$ in (10). The last term $\ln b N_2$ or $\ln b N_2{}^*$ or $N_2{}^*$, which has been obtained by researching into the orders of magnitude of quantities (or the limiting processes) and seems to have been overlooked hitherto, is seen to be essential to the existence of an exactly horizontal part of an isotherm in the limit $N \to \infty$. As one compresses a gas at constant temperature, the series $\sum_{l=1}^{\infty} b_l z^l$ (at the same time, also $\sum_{l=1}^{\infty} l b_l z^l$) becomes singular at the point $v = v_s$, $z = z_s$ that is the starting-point of condensation, and then instantly the last term $\ln b N_2{}^*$ in (18) (at the same time, also $N_2{}^*$ in (32)) appears, that is to say, one "large" cluster consisting of $N_2{}^*$ molecules, which may correspond to the liquid phase in equilibrium with the vapour, appears and is described by the expression

$$\ln b_L \rightleftharpoons L \sum_{k=1}^{k<L} \beta_k y_0{}^k - L \ln y_0 \tag{33}$$

or

$$\ln b_L \equiv L \sum_{k=1}^{k \ll L} \beta_k y_s{}^k - N_2{}'^* \ln y_s + \ln \beta N_3{}^*. \quad (34)$$

Here it should be emphasized that, when the series for the gaseous phase becomes singular, the theory of condensing systems does *not* break down by the divergence of the series, but a new term or series appears and plays the leading role, in place of the old series which has already become singular.

The two-phase separation that is the appearance of liquid in the condensation range, which we have derived from the last term as stated above, has not been derived by the Kahn-Uhlenbeck-Born-Fuchs[3, 4] method of treatment.

Further researches for the last term may lead to the knowledge of liquid phase. In fact, we have a method to evaluate the volume dependence of $\ln b_L$, and on the other hand we may consider b_l ($l \ll N$) to be independent of volume, and hence we have some possibility that we can derive the properties of liquid phase and the singularity corresponding to

solidification automatically from the partition function expanded by the Ursell method.

It should be remarked that our method of treatment of the Ursell expansion has analogy with the theories and criticisms of Bose condensation. [Comp. F. London's[5] first term. On the other hand, comp. G. Schubert's[6] and R. B. Dingle's[7] criticisms of the saddle-point method.]

1) H. D. Ursell, Proc. Cambr. Phil. Soc. **23** (1927), 685.

2) J. E. Mayer, et al., J. Chem. Phys. **5** (1937) 67, 74, **6** (1938), 87, 10.

3) B. Kahn and G. E. Uhlenbeck, Physica **5** (1938), 399.

4) M. Born and K. Fuchs, Proc. Roy. Soc. A **166** (1938), 391.

5) F. London, Phys. Rev. **54** (1938), 947.

6) G. Schubert, Z. Naturforsch. **1** (1946), 113; **2** (1947), 250.

7) R. B. Dingle, Proc. Cambr. Phil. Soc. **45** (1949), 275.

The Formulation of Quantum Mechanics in terms of Ensemble in Phase Space[1]

Takehiko TAKABAYASI

Physical Institute, Nagoya University

(Received November 16, 1953)

The formulation of non-relativistic quantum mechanics in terms of ensemble in phase space is established by clarifying the subsidiary conditions for the phase space ensemble to represent a pure state, and thereby the equivalent correspondence between this formulation and the alternative formulation in terms of quantum potential previously developed is exhibited.

§ 1. Introduction and summary

The ordinary formulation of quantum mechanics, as established by the fusion of Heisenberg's matrix mechanics and Schrödinger's wave mechanics, is certainly the most fundamental and powerful one, having its own ' picture ' in a broad sense* essentially non-classical.   Nevertheless we may consider another consistent formulation of quantum mechanics with its associated picture, for instance, path integral formulation by Feynman[3].   Generally such a new formulation and picture would reveal new aspects of physical and mathematical construction of quantum mechanics, and might serve to suggest new clues to future progress of quantum theory itself**, apart from its usefulness for practical applications to specified class of problems.

From such viewpoint we have examined in detail a certain formulation of quantum mechanics in previous papers[4][5] (see § 5 (a)):   The method is based on the transformation of customary Schrödinger equation into simultaneous equations for the phase and amplitude of the wave function, which are found to be of the form of Hamilton-Jacobi-like equation or Euler's equation of motion for velocity potential, and the equation of continuity. According to this expression we have the representation of quantum mechanical motion in terms of an ensemble of trajectories in configuration space subject to some additional force (so-called ' quantum force '), or equivalently in terms of an irrotational flow of perfect fluid with peculiar internal stress (' quantum stress ').   We shall call this method the method of the *configuration space ensemble* (abbreviated as cs. en.).

Now the Schrödinger equation can be transformed into a form describable in classical

---

*) Extending " the meaning of the word ' picture' to include any way of looking at the fundamental laws which makes their self-consistency obvious ", according to Dirac[2].

**) In this paper, however, we shall not try any such suggestion, confined merely in the reformulation of the present quantum theory.

languages in still another way :   That is, we can transform the Schrödinger equation into Liouville-like equation for a distribution function in phase space which is produced as a certain fourier transform of a bilinear form of the wave function, leading to the picture of certain Markoff-like process of an ensemble in phase space for quantum-mechanical motion. This method, which we will refer to as the method of the *phase space ensemble* (abbreviated as ps. en.), was initiated by Wigner[6] and later by Moyal[7].   The purpose of the present paper is to develop this method into a consistent formulation of quantum mechanics by establishing the *subsidiary conditions* for a ps. en. to represent a *pure state*, and also to prove thereby the equivalence of this formulation with the cs. en. formulation formerly mentioned.

In the phase space formulation the knowledge involved in the phase of the original wave function is reflected upon the momentum distribution in such a manner that the phase space distribution function (abbreviated as ps. df.) implies the representation of a state symmetrical in coordinates and momenta.   But the manifold of the ps. df. covers wider possibilities than that of the original wave functions.   Now, according to our prescription, a *mixing* of states corresponds to a *superposition* with positive coefficients of relevant distribution functions which as well satisfies the same Liouville-like equation, because the latter is linear in the df.   Accordingly a ps. df. in general would be the representative of a mixed state, in so far as it satisfies certain ' positivity condition '.   Thus in this formulation of quantum mechanics it is an essential problem to obtain the subsidiary conditions[*] that a ps. df. should particularly correspond to a pure state.   We explicitly obtain these conditions, which must be of some non-linear relations (§ 4).   This is made tractable by first replacing the usual pure state condition for the density matrix (4.2) by *local relations* (4.4).   Transforming the latter we acquire the pure state conditions on the ps. df., which consist of the condition of irrotationality of mean momentum field, (4.19), and the condition (4.6) which we call the ' quantum condition '.   The latter will further be transformed into a series of relations between distribution moments in respect to momentum components of successively higher orders.

Now, in virtue of these pure state conditions, we can prove the equivalence and correspondence between the cs. en. formulation and the ps. en. formulation (§ 5).   For instance, the quantum potential in the former may be looked upon as an apparent force appearing as a result of ' projecting ' the ps. en. onto the configuration space.

The cs. en. formulation and ps. en. formulation, though they are equivalent and transmutable to each other, are of very different characters.   The effect of quantum fluctuations is represented with fluctuations of continuous trajectories[5] due to quantum potential in the former, while in the latter with Markoff-like transitions, the properties of which we shall examine in detail (§ 3).   The ps. en. seems to be one step superior to the cs. en. in that it correctly yields quantum-mechanical expectation values as the mean values over the ensemble for wider class of dynamical quantities, yet it must be emphasized that

---

[*]   Moyal[7] unnoticed the presence of these conditions, while other authors have been unable to obtain their expression.

it cannot do so for *all* hermitian quantities and that we must in return allow of *negative probabilities* (§ 2).

The ps. en. formulation is formally consistent within its limited range of applicability and accompanied with the picture working along classical lines, but we cannot take the picture too realistically, just as in the case of the cs. en. formulation[4]. Instead, these formulations provide *concrete analyses as to the degree in which the statistical properties of quantum mechanics can be understood along any statistical scheme based on some hidden variables.*

The ps. df. is a *real* quantity produced as a *bilinear* form of the wave function, as is needed for the representation in terms of it to have classical pictures, but that would just mean greater complications in mathematical treatments usually. It is well known that the method of ps. df. is useful for the treatment of quantum statistical mechanics[6)8]. We shall, however, show how this method of ps. en. can effectively be applied to pure state problems for a few elementary examples (§ 6).

In the last section we consider the positivity condition and also express the pure state condition in an alternative form.

§ 2.   Phase space distributions and mean values

For simplicity, we shall confine ourselves to the simplest case of a single non-relativistic particle without spin throughout*. A quantum-mechanical mixed state can be specified by a density matrix[a] $\underline{\rho}$, or $\langle x|\underline{\rho}|x'\rangle \equiv \rho(x, x')$ in coordinate representation, which must satisfy the conditions to be

hermitian : $\qquad\qquad\qquad\qquad \rho(x, x') = \rho^*(x', x),$ $\qquad\qquad\qquad$ (2·1)

normalizable : $\qquad\qquad\qquad \mathrm{Sp}\,\underline{\rho} = \int_{-\infty}^{\infty} \rho(x, x)\,dx = 1,$ $\qquad\qquad$ (2·2)

and positive definite : $\qquad\quad \mathrm{Sp}\,(\underline{\rho}\,\underline{A}^2) \geq 0,$

$\qquad\qquad\qquad\qquad\quad$ for any hermitian operator $\underline{A}.$ $\qquad\qquad$ (2·3)**

Conversely, any function $\rho(x, x')$ satisfying these conditions can be expanded in a form

$$\rho(x, x') = \sum_n w_n \psi_n(x) \psi_n^*(x'), \quad\Big\}$$
$$\text{with} \qquad w_n \geq 0, \quad \sum_n w_n = 1, \qquad\qquad (2\cdot4)$$

where $\psi_n(x)$ and w_n mean eigen-functions and eigen-values of $\underline{\rho}$ respectively, which fact indicates that the $\rho(x, x')$ corresponds to a mixture of pure states of wave functions $\psi_n(x)$ with respective weights w_n.

The density matrix $\underline{\rho}$ gives the quantum-mechanical expectation value of any quantity

---

*)   The region of applicability of our formulation is rather limited. When a vector potential is acting, the Markoff-like picture to be stated in § 3 requires certain extension.

**)   In this paper we indicate an abstract operator by attaching an underline.

$\underset{=}{A}$ by

$$\langle \underset{=}{A} \rangle_{\mathrm{qu}} = \mathrm{Sp}(\rho \ \underset{=}{A}). \tag{2.5}$$

Now, from ρ we define, according to Wigner[6], a function $f(x, p)$ by the transformation,

$$f(x, p) = \frac{1}{(2\pi\hbar)^3} \int \rho\left(x - \frac{y}{2}, \ x + \frac{y}{2}\right) e^{ipy/\hbar} \, dy, \tag{2.6}*$$

$$= \frac{1}{(2\pi\hbar)^3} \rho\left(x - \frac{\hbar}{2i}\nabla_p, \ x + \frac{\hbar}{2i}\nabla_p\right) \delta(p),$$

$$= (\pi\hbar)^{-3} \int \langle p|x'\rangle \langle x'|\rho|x''\rangle \langle x''|p\rangle \delta(x' + x'' - 2x) \, dx' dx'',$$

and will use the latter for the specification of the mixed state. The function can also be written in the *reciprocal* form :

$$f(x, p) = \frac{1}{(2\pi\hbar)^3} \int \rho\left(p - \frac{q}{2}, \ p + \frac{q}{2}\right) e^{-iqx/\hbar} dq. \tag{2.7}*$$

where $\rho(p, p')$ is the momentum representation of the density matrix. Eqs. (2.6) and (2.7) show that the function $f(x, p)$ is the fourier transform of $\rho(x, x')$ or $\rho(p, p')$ along its antidiagonal ; in other words, if we consider the density matrix to be a function of 'mean coordinate' and 'relative coordinate', as

$$\bar{\rho}(x, y) = \rho(x - y/2, \ x + y/2),$$

f is the fourier transform of $\bar{\rho}$ in respect to the relative coordinate.

The function $f(x, p)$ should be, corresponding to (2.1),

$$f(x, p) = \text{real, (though not necessarily positive)}, \tag{2.8}$$

and satisfy, corresponding to (2.2),

$$\int f(x, p) \, dx dp = 1. \tag{2.9}$$

We may thus imagine, corresponding to a mixed state, an 'ensemble' with the 'probability distribution' in *phase space*, $f(x, p)$, though we must then allow of negative probabilities. The function $f(x, p)$, which we call the phase space distribution function (ps. df.), can be regarded as the *representative of state symmetrical in coordinate and momentum*. From (2.6) or (2.7) we get

$$\begin{cases} P(x) = \int f(x, p) \, dp = \langle x|\rho|x\rangle, \tag{2.10} \\ Q(p) = \int f(x, p) \, dx = \langle p|\rho|p\rangle; \tag{2.11} \end{cases}$$

that is, our ps. df. $f(x, p)$ leads to the positional distribution and momentum distribu-

---

*) In this paper integral usually means a definite integral over whole coordinate space or whole momentum space. In such case we shall hereafter omit to note the boundaries $\pm\infty$.

tion of quantum-mechanical particle correctly. Accordingly, we have further

$$\langle \underline{A} \rangle_{qu} = \mathrm{Sp}\,(\rho \underline{A}(x,\ p)) = \int A(x,\ p) f(x,\ p)\, dx dp, \qquad (2\cdot 12\mathrm{a})$$

provided $\underline{A}(x,\ p)$ belongs to the quantities that separate as

$$\underline{A}(x,\ p) = \underline{A}_1(x) + \underline{A}_2(p). \qquad (2\cdot 12\mathrm{b})$$

This means that our phase space ensemble (ps. en.) correctly gives the quantum mechanical expectation values as the average values over the ensemble for any quantity of type $(2\cdot 12\mathrm{b})$[6].

However, for a more general function $A(x,\ p)$ involving cross terms of x and p there exist ambiguities in the definition of the corresponding quantum-mechanical function $\underline{A}(x,\ p)$ of non-commutable operators $\underline{x}$ and $\underline{p}$, so that in such a case our ps. df. may give the mean value correctly only for a special quantity $\underline{A}^{(W)}(x,\ p)$ among the quantities $\underline{A}(x,\ p)$ corresponding to the same c-number function $A(x,\ p)$. Such $\underline{A}^{(W)}(x,\ p)$ is the same with that defined by Weyl's[9] procedure, i.e.,

$$\underline{A}^{(W)}(x,\ p) = \int a(\sigma,\ \tau) e^{i(\sigma \underline{x} + \tau \underline{p})}\, d\sigma\, d\tau, \qquad (2\cdot 13)$$

where $a(\sigma,\ \tau)$ is the classical fourier coefficient of $A(x,\ p)$:

$$A(x,\ p) = \int a(\sigma,\ \tau) e^{i(\sigma x + \tau p)}\, d\sigma\, d\tau. \qquad (2\cdot 14)$$

The above mentioned validity of the relation,

$$\langle \underline{A}^{(W)}(x,\ p) \rangle_{qu} = \int A(x,\ p) f(x,\ p)\, dx\, dp, \qquad (2\cdot 15)$$

is clear from another expression of the ps. df.:

$$f(x,\ p) = (2\pi)^{-6} \int \mathrm{Sp}\,(\rho e^{i(\sigma \underline{x} + \tau \underline{p})}) e^{-i(\sigma x + \tau p)}\, d\sigma\, d\tau. \qquad (2\cdot 16)[7]$$

Thus it may readily be found that the ps. df. does not give mean values correctly for quantities such as, e.g., the commutator $[\underline{p}_i,\ \underline{x}_i]$, the square of the energy H^2 (except for the case of free particle), or the magnitude* of angular momentum l^2. In other words our ps. df. does not reflect quantum-mechanical probability distributions correctly for quantities such as $\underline{H}$ or $\underline{l}_i$ (a component of angular momentum), in contrast to the case of $\underline{A}_1(x)$ or $\underline{A}_2(p)$.

By the way, our ps. df. can also be written as

$$f(x,\ p) = \exp(\hbar/2i \cdot \nabla_p \nabla_x)\, g(x,\ p), \qquad (2\cdot 17)$$

in terms of

$$g(x,\ p) = \langle p|\rho|x\rangle \langle x|p\rangle. \qquad (2\cdot 18)$$

Here the latter distribution function $g(x,\ p)$, though not real, gives the mean values correctly for the 'well-ordered' functions[10] of the form, $\underline{A}(x,\ p) = \sum_{n,\ m} a_{nm}\, \underline{x}^n\, \underline{p}^m$.

---

*) For the components of angular momentum, the ps. df. yields expectation values correctly.

Now we must consider the condition $(2\cdot3)$. Taking a function of x alone, or of p alone, $(2\cdot3)$ requires

$$\begin{cases} \mathrm{Sp}(\rho\, A_1(x)^2) = \int A_1(x)^2 P(x)\, dx \geq 0, \\ \mathrm{Sp}(\rho\, A_2(p)^2) = \int A_2(p)^2 Q(p)\, dp \geq 0. \end{cases}$$

These relations hold for any real function $A_1(x)$ or $A_2(p)$, so that we must have

$$P(x) \geq 0 \quad \text{and} \quad Q(p) \geq 0. \tag{2·19}$$

This is a necessary condition but not sufficient for $(2\cdot3)$ (see § 7).

We could represent a general quantum-mechanical mixed state with a phase space ensemble, whose distribution function $f(x, p)$ is not itself necessarily positive everywhere, though required to produce a positive configuration-space df. $P(x)$ and positive momentum-space df. $Q(p)$. Conversely, any ps. df. $f(x, p)$ satisfying $(2\cdot8)$, $(2\cdot9)$, and certain 'positivity condition' including $(2\cdot19)$, is a permissible one corresponding to a mixed state. The non-positive-definiteness of our general ps. df., which stems from the same property of the ps. df. for pure states, discloses the physically unreal nature of our ps. en. It is a general characteristic of quantum-mechanical probability distributions for non-commuting quantities that they cannot be derived from a single statistical ensemble based on hidden variables, at least without admitting of negative probabilities.

When we are given the configuration and momentum distribution functions $P(x)$ and $Q(p)$ independently, except for the normalization condition,

$$\int P(x)\, dx = \int Q(p)\, dp = 1, \tag{2·20}$$

we can construct many ps. dfs. which are compatible with those given $P(x)$ and $Q(p)$. Indeed the ps. df.

$$f_0(x,p) = P(x) Q(p) \tag{2·21}$$

is clearly such one, and moreover any ps. df.

$$f(x, p) = P(x) Q(p) + f_1(x,p) \tag{2·22}$$

in which f_1 is an arbitrary function staisfying

$$\int f_1(x, p)\, dp = \int f_1(x, p)\, dx = 0, \tag{2·23}$$

also leads to the given $P(x)$ and $Q(p)$. Such ps. df., $(2\cdot21)$ or $(2\cdot22)$, is a permissible one, corresponding generally to a mixed state, provided it satisfies the positivity condition. The ps. df. that factorizes as $(2\cdot21)$ is a particular one having no correlation at all between particle position and momentum*[)]. In this connection, it is further to be

---

*) Such property cannot generally be conserved with the passage of time in classical as well as in quantum ps. en. In the former, however, this property persists in the special case of stationary canonical ensemble, while in the latter it is not so. This is due to the difference of the 'Liouville' equations for ps. dfs. in both cases (see § 3).

remarked that there can exist many distribution functions in phase space other than $f(x, p)$, which satisfy $(2 \cdot 8)$ and $(2 \cdot 9)$ and give expectation values correctly for any quantity of type $(2 \cdot 12b)$. Examples are

$$f_a(x, p) = \cos(\hbar/2 \cdot \nabla_p \nabla_x) f(x, p),$$

and

$$f_b(x, p) = \mathrm{Re}[g(x, p)].$$

We have considered the df. defined in $x - p$ space, but it may also be possible to introduce a distribution as a function of another variables ξ and η, where ξ or η represents each a complete set of commuting observables, in such a way that it correctly gives the quantum-mechanical probability distributions for quantities of type $A_1(\xi)$ and $A_2(\eta)$; this distribution*, however, differs from the original one, $f(x, p)$. In this sense the ps. en. has not the meaning invariant to such pairs of unitary transformations[11] $(\underline{x}, \underline{p} \rightarrow \underline{\xi}, \underline{\eta})$.

§ 3. Time development of the distribution

(a) Equation of motion and the transition in momentum

We shall now represent the temporal change of state in terms of ps. en. The equation of motion for $\underline{\rho}$,

$$i\hbar \, d\rho/dt = \underline{H}\rho - \rho \underline{H}, \qquad (3 \cdot 1)$$

is written, in x-representation, as

$$i\hbar \, \partial\rho/\partial t = -\hbar^2/2m \cdot (\varDelta_x - \varDelta_{x'})\rho + (V(x) - V(x'))\rho, \qquad (3 \cdot 2)$$

since we take hamiltonian, $H = p^2/2m + V(x)$, corresponding to a particle in a scalar potential $V(x)$. For the intermediary function

$$\bar{\rho}(x, y) = \rho(x - y/2, x + y/2) = \int f(x, p) e^{-ipy/\hbar} \, dp, \qquad (3 \cdot 3)$$

$(3 \cdot 2)$ is written as

$$i\hbar \, \partial\bar{\rho}/\partial t = \hbar^2/m \cdot \nabla_x \nabla_y \, \bar{\rho} + \{V(x - y/2) - V(x + y/2)\}\bar{\rho}, \qquad (3 \cdot 4)$$

which, by fourier transformation, leads to the equation of motion for our df.:

$$\partial f/\partial t + p/m \cdot \nabla f = \Lambda[f], \qquad (3 \cdot 5)$$

with

$$\Lambda[f] = \int J(x, p - p') f(x, p') \, dp', \qquad (3 \cdot 6)$$

$$J(x, p) = -\frac{2}{\hbar} \frac{1}{(2\pi\hbar)^3} \int V\left(x + \frac{y}{2}\right) \sin \frac{py}{\hbar} \, dy \qquad (3 \cdot 7)$$

---

*) It is given by

$$f(\xi, \eta) = (\pi\hbar)^{-3} \int \langle \eta | \xi' \rangle \langle \xi' | \rho | \xi'' \rangle \langle \xi'' | \eta \rangle \delta(\xi' + \xi'' - 2\xi) d\xi' \, d\xi''.$$

$$= -\frac{2^4}{\hbar}\, \mathrm{Im}\,[\,V(2p)\,e^{-2ip\cdot x/\hbar}],\qquad (3\cdot8)$$

where

$$V(p)=\frac{1}{(2\pi\hbar)^3}\int V(x)\,e^{ipx/\hbar}\,dx.\qquad (3\cdot9)$$

The appearance of the fourier component, $V(2p)$, may more easily be understood by starting from the momentum representation of $(3\cdot1)$,

$$i\hbar\,\frac{\partial\rho(p,p')}{\partial t}=\frac{1}{2m}(p^2-p'^2)\rho(p,p')$$

$$+\int(V(p-p'')\rho(p'',\,p')-\rho(p,\,p'')V(p''-p'))dp'',$$

and transforming it by use of $(2\cdot7)$.

By expansion J may also be written as

$$J(x,\,p)=\sum_{n}{}'\sum_{n_1+n_2+n_3=n}\frac{(-\mathcal{E})^{(n-1)/2}}{n_1!\;n_2!\;n_3!}\,\frac{\partial^n V}{\partial x_1^{n_1}\partial x_2^{n_2}\partial x_3^{n_3}}\,\frac{\partial^n\delta(p)}{\partial p_1^{n_1}\partial p_2^{n_2}\partial p_3^{n_3}},\qquad (3\cdot10)$$

where $\mathcal{E}\equiv\hbar^2/4$, and $\sum_{n}{}'$ means the summation over positive odd integers. Using this form, the integral operator A is expressed as an infinite series of differential operators[6)7)]:

$$A[f]=\sum_{n}{}'\sum_{n_1+n_2+n_3=n}\frac{(-\mathcal{E})^{(n-1)/2}}{n_1!\;n_2!\;n_3!}\,\frac{\partial^n V}{\partial x_1^{n_1}\partial x_2^{n_2}\partial x_3^{n_3}}\,\frac{\partial^n f}{\partial p_1^{n_1}\partial p_2^{n_2}\partial p_3^{n_3}},\qquad (3\cdot11)$$

$$=2/\hbar\cdot\sin\!\left(\frac{\hbar}{2}\nabla_x\,\nabla_p\right)V(x)f(x,\,p),\qquad (3\cdot12)$$

with the understanding that ∇_x in $(3\cdot12)$ operates on $V(x)$ alone. Expressions $(3\cdot10)$ and $(3\cdot11)$ may be regarded as expansions in ascending powers of $\mathcal{E}$, of which first terms are

$$J^{(0)}(x,\,p)=\nabla V\cdot\nabla_p\delta(p),\qquad A^{(0)}[f]=\nabla V\cdot\nabla_p f.\qquad (3\cdot13)$$

If $V(x)$ is a polynomial in or below second power, only the first terms $(3\cdot13)$ in the

---

*) To next page: The circumstance that in case of potential quadratic (including below quadratic) in x, the ps. en. moves purely classically is closely connected with the fact that in this case, speaking with the language of matrix mechanics, the Hamilton's or Newton's *equation of motion* for q-number coordinates is *linear*[12)]. Generally in quantum mechanics formally the same equation of motion as in classical theory holds for q-number quantities, but, if this equation of motion be linear, there does not appear any *product of q-numbers*, the rule on which specifies the essential difference of quantum mechanics from classical theory, and therefore each matrix element changes like classical quantity. In this case the difference of quantum mechanics from classical theory can only present itself in the definition of the initial conditions which is represented by the commutation relation in matrix mechanics, and by the subsidiary conditions for pure state (to be stated in § 4) in our formulation. It is to be noted that the above case includes a generalized *forced* ' oscillator ' with the hamiltonian, $H=\alpha(t)\underline{p}^2+\beta(t)\underline{x}^2+\gamma(t)\underline{x}$.

:

series survive, and then $(3\cdot5)$ reduces to the form identical with the Liouville equation for the classical df. Therefore in such a case the time development of our ps. en. can be regarded as is produced by the process that each point of the ensemble moves along its purely classical continuous trajectory, in precisely the same manner as in the classical statistical ensemble.*

Such a picture, however, fails in general cases, where $A[f]$ is an integral operator in momentum space. Yet, looking upon the second and higher order terms in the expansion $(3\cdot11)$ as quantum-mechanical corrections, we call $(3\cdot5)$ 'quantum-mechanical Liouville equation'. The formal interpretation of this equation leads to the following *stochastic* picture for the time development of the ensemble : The coordinate x of each particle of the ensemble changes continuously with the velocity p/m, while the value of its momentum jumps with a 'transition probability' J ($J(x, p)dp$ meaning the probability with which the momentum jumps in unit time by an amount $p \sim p+dp$ at point x).

Now this stochastic picture has further following features : (i) J is an odd function of p ; therefore it takes negative as well as positive values, and also

$$\int J(x, p')dp' = 0, \qquad (3\cdot14)$$

that is, this transition probability is not normalizable, though $(3\cdot14)$ serves to ensure the distribution probability :

$$\frac{d}{dt}\int f \, dx \, dp = 0. \qquad (3\cdot15)$$

(ii) The transition probability depends upon the amount of jump alone, irrespective of the value of momentum before or after the jump. (iii) The external field $V'(x)$ acts so as to induce indeterministic transitions in particle momentum, but the transition probability J itself is perfectly determined by this potential. (iv) As is seen in $(3\cdot8)$ $J(x, p)$ is of a form of a perfect sinusoidal wave in x-space with no damping in far away irrespective of the form of $V'(x)$. The amplitude of this wave is the fourier component** $V'(2p)$; while its wave-length is $\hbar/2p$, so that for larger jumps of momentum the x-space oscillation of J is more rapid, having therefore less effects when we consider the space average over the ensemble. (v) Since

$$\int pJ(x, p)dp = -\nabla V, \qquad (3\cdot16)$$

the rate of average change of momentum agrees with the classical value. In each individual case, however, particle fluctuates, performing jumps in momentum, which may well be far greater than the mean value, and may thus, for instance, penetrate a potential barrier. (vi) The various 'transition moments' of momentum components are found to be

---

**) The circumstance that the probablity of momentum transition from p_0 to p is determined by $V'(2(p-p_0))$ in our stochastic picture is somewhat similar to the usual perturbation-theoretical result of quantum mechanics, where the probability of transition from the state of momentum p_0 to that of p is proportional to $|V(p-p_0)|^2$.

$$m_{n_1 n_2 n_3} \equiv \int p_1^{n_1} p_2^{n_2} p_3^{n_3} \, J(x, \, p) dp$$

$$= \begin{cases} -(-\varepsilon)^{(n-1)/2} \dfrac{\partial^n V}{\partial x_1^{n_1} \partial x_2^{n_2} \partial x_3^{n_3}}, & \text{if } n = \sum_i n_i = \text{odd,} \\[2mm] 0, & \text{if } n = \text{even,} \end{cases} \quad (3 \cdot 17)$$

including $(3 \cdot 14)$ and $(3 \cdot 16)$ as special cases. On account of the odd character of J, transition moments of even orders vanish while those of odd orders survive, in contrast to the usual Brownian processes. In the latter, the transition moments in and above third orders are assumed to vanish, resulting in the differential equation of Fokker-Planck-Kramers[13] type for the df.; the diffusion takes place in a manner essentially determined by the second order moment, the distribution always diffusing monotonously and irreversibly. On the other hand in our case the moments of third and higher odd orders give rise to a quite different type of ‘diffusion’[*)**)***]. The non-vanishing of third or higher order moments means that in our process the probability that the value of momentum changes by a finite amount in a small time interval cannot be regarded as small.

We have thus obtained a stochastic picture of distinctive features for a quantum-mechanical change of state. Though this picture cannot be taken as a real one, for instance, on account of (i), it may be said to represent quantum fluctuations in a picturesque manner.

(b) Transition probability for a finite time interval

Our df. develops with time according to the linear integro-differential equation $(3 \cdot 5)$ which is of the first degree in t, and therefore, given the initial distribution $f(x_0, p_0, t_0)$, later distributions will be uniquely determined. The time development of f may thus be written in an integral form,

$$f(x, \, p, \, t) = \int T(xpt | x_0 p_0 t_0) f(x_0, \, p_0, \, t_0) dx_0 \, dp_0, \quad (3 \cdot 18)$$

where the kernel $T(xpt | x_0 p_0 t_0)$ embodies the temporal development law of the distribution independently of the initial condition, and implies the *transition probability* in phase-space for finite time interval, i.e., the distribution at t conditional in x_0, p_0 at t_0.

Naturally this function is closely connected with the kernel (propagation function) for the time development of the wave function $\psi(x, \, t)$, which is written as

$$\psi(x, \, t) = \int K(xt | x_0 t_0) \psi(x_0 t_0) dx_0. \quad (3 \cdot 19)$$

---

*) Cf. § 6 (a) and appendix A.

**) Such a result corresponds to the circumstance that the wave function (the probability *amplitude*) satisfies a diffusion-type differential equation with imaginary time coefficient, and therefore, when we transform this equation to that for a real quantity interpretable as a probability *distribution* by an iteration procedure, the latter equation can no longer be of a diffusion type.

***) In conventional Brownian processes, furthermore, there acts a frictional force which is proportional to particle velocity and makes it tend to the equilibrium distribution. In our process, however, such effects do not occur and it is impossible that any distribution should converge with the passage of time to some equilibrium distribution that would correspond to a stationary state [see § 6 (b) iii)].

This kernel $K(xt|x_0 t_0)$ is nothing but the transformation function which transforms the representation making x_{t_0} (particle position at t_0) diagonal to the one making x_t (position at t) diagonal, i.e.,

$$K(xt|x_0 t_0) = \langle x_t' | x_{t_0}' \rangle, \qquad \text{with } x = x_t', \ x_0 = x_{t_0}', \text{ numerically.} \tag{3.20}$$

Hence it satisfies, as is well known, the iteration law :

$$K(xt|x_0 t_0) = \int K(xt|x_1 t_1) K(x_1 t_1|x_0 t_0) dx_1, \tag{3.21}$$

and the unitarity condition :

$$\int K(xt|x_0 t_0) K(x_0 t_0|x't) dx_0 = \delta(x - x'), \tag{3.22}$$

and so

$$K(xt_0|x_0 t_0) = \delta(x - x_0). \tag{3.23}$$

As $K(xt|x_0 t_0)$ also satisfies the Schrödinger equation, it can be determined by solving that equation under the initial condition (3.23). For a conservative system K may be written as

$$K(xt|x_0 t_0) = \langle x|e^{-iH(t-t_0)/\hbar}|x_0 \rangle, \tag{3.24}$$

which is a function of $t - t_0$ alone in respect to the time, and satisfies

$$K(xt|x_0 t_0) = K^*(x_0 t_0|xt). \tag{3.25}$$

Now, corresponding to (3.19), the kernel for the time development of the density matrix $\rho(x, x', t)$ is given by

$$\int K(xt|x_0 t_0) K^*(x't|x_0' t_0) dx_0 dx_0', \tag{3.26}$$

and accordingly the relation of our transformation function T with K is found to be

$$T(xpt|x_0 p_0 t_0) = \frac{1}{(2\pi\hbar)^3} \int K\left(x - \frac{y}{2}, t \,\middle|\, x_0 - \frac{y_0}{2}, t_0 \right) \times$$

$$\times K^*\left(x + \frac{y}{2}, t \,\middle|\, x_0 + \frac{y_0}{2}, t_0 \right) e^{(i/\hbar)(py - p_0 y_0)} \, dy \, dy_0. \tag{3.27}$$

From the properties of K, (3.21)–(3.25), we can find the corresponding properties of T. First, T is real, satisfies the iteration law :

$$T(x \, p \, t|x_0 \, p_0 \, t_0) = \int T(x \, p \, t|x_1 \, p_1 \, t_1) T(x_1 \, p_1 \, t_1|x_0 \, p_0 \, t_0) dx_1 \, dp_1, \tag{3.28}$$

and the unitarity :

$$\int T(x \, p \, t|x_0 \, p_0 \, t_0) T(x_0 \, p_0 \, t_0|x' \, p' \, t) dx_0 \, dp_0 = \delta(x - x') \, \delta(p - p'), \tag{3.29}$$

and so we also have

$$T(xpt_0|x_0 p_0 t_0) = \delta(x - x_0) \, \delta(p - p_0). \tag{3.30}$$

The T-function itself satisfies the quantum Liouville equation :

$$\frac{\partial T(xpt|x_0 p_0 t_0)}{\partial t} + \frac{p}{m} \nabla T = \int J(x, p - p') T(x \, p' \, t|x_0 \, p_0 \, t_0) \, dp', \tag{3.31}$$

and we may determine T by solving $(3 \cdot 31)$ under the initial condition $(3 \cdot 30)$. For a conservative system, $T(x\,p\,t | x_0\,p_0\,t_0)$ depends on $t - t_0$ alone in respect to the time, and is symmetric :

$$T(x\,p\,t | x_0\,p_0\,t_0) = T(x_0\,p_0\,t_0 | x\,p\,t), \qquad (3 \cdot 32)$$

and normalized :

$$\int T(x\,p\,t | x_0\,p_0\,t_0)\,dx\,dp = 1. \qquad (3 \cdot 33)$$

Eq. $(3 \cdot 28)$ with $(3 \cdot 30)$ and $(3 \cdot 33)$ shows that our stochastic process is a sort of *Markoff process* in phase-space with the transition probability T, though we must allow of negative probabilities, since T as well as f is not necessarily positive. The law of composition $(3 \cdot 28)$ may be regarded as the integral equation (" Smoluchowski equation ") for the transition probability function T, which may be reduced to the 'differential form' $(3 \cdot 31)$ with $(3 \cdot 10)$. The latter equation, however, is of a type quite different from the Fokker-Kramers-Kolmogoroff equation for usual Brownian processes, since, in our process, the differential coefficients of V in and above third degree, in so far as they do not vanish, play the roles of higher transition moments for momentum components.

Eq. (3.31) with (3.30) is written as

$$T(xpt | x_0 p_0 t_0) = \delta(x - x_0)\ \delta(p - p_0) - \int_{t_0}^{t} p/m \cdot \nabla T$$
$$+ \int_{t_0}^{t} dt \int J(x,\ p - p')\ T(xp't | x_0 p_0 t_0)\,dp', \qquad (3.34)$$

which may be solved by iteration as a power series in $t - t_0$:

$$T(xpt | x_0 p_0 t_0) = \delta(x - x_0)\delta(p - p_0) + (t - t_0)\,T_1 + \frac{(t - t_0)^2}{2}\,T_2 + \cdots,$$

with

$$T_1(xp | x_0 p_0) = -\frac{p}{m}\,\delta(p - p_0)\nabla_x \delta(x - x_0) + \delta(x - x_0)\,J(x, p - p_0),$$

$$T_2(xp | x_0 p_0) = -\frac{1}{m^2}\,\delta(p - p_0)\,(p\nabla)\,(p_0\nabla_0)\delta(x - x_0)$$

$$\frac{p + p_0}{m}\nabla_x\delta(x - x_0)\,J(x, p - p_0) + \delta(x - x_0)\int J(x, p - p')\,J(x, p' - p_0)\,dp'.$$

When the potential $V(x)$ is, in particular, a polynomial in or below second power, T is the solution of the classical Liouville equation with the initial condition (3.30), and hence is given by

$$T(xpt | x_0 p_0 t_0) = \delta(x - x_t(x_0 p_0 t_0))\delta(p - p_t(x_0 p_0 t_0)), \qquad (3.35)$$

where $(x_t,\ p_t)$ is the solution of the classical Newtonian equation of motion. For example, for a free particle we have

$$T(xpt | x_0 p_0 t_0) = \delta(p - p_0)\delta(x - x_0 - p_0/m \cdot (t - t_0)), \qquad (3.36)$$

which is quite different from the kernel for the usual diffusion, while the kernel for probability amplitude is, as is well known, of diffusion type with imaginary time coefficient :

$$K(xt | x_0 t_0) = \left(\frac{m}{2\pi\hbar i(t - t_0)}\right)^{3/2} \exp\left(\frac{im}{2\hbar}\frac{(x - x_0)^2}{t - t_0}\right). \qquad (3.37)$$

But the fourier transformation of (3.36), taking account of the relation (3.27), gives

$$K\left(x-\frac{y}{2}, t\,\bigg|\,x_0-\frac{y_0}{2}, t_0\right)K^*\left(x+\frac{y}{2}, t\,\bigg|\,x_0+\frac{y_0}{2}, t_0\right)$$

$$=-\left(\frac{m}{2\pi\hbar(t-t_0)}\right)^3\exp\left(-\frac{im}{\hbar}\,\frac{(x-x_0)(y-y_0)}{t-t_0}\right), \tag{3.38}$$

from which we can derive (3.37).

§ 4. Subsidiary conditions for pure state

(a)   As was stated in § 2, generally a ps. df. $f(x, p)$, in so far as it satisfies the positivity condition, corresponds to a mixed state, and the procedure of mixing of states corresponds to a superposition (with positive coefficients) of relevant distribution functions. Therefore, various relations thus far stated which are valid for any mixed states (including pure states as their special cases) must be linear ones (except for the positivity condition). In fact, the expression of mean values (2.12), and the time development equation (3.5) or (3.18) are all linear in f. Therefore the subsidiary condition that must be imposed upon a df. in order that the df. should in particular correspond to a pure state, must be some non-linear relation. It must be further of such character as to restrict the functional space of the df. which is one real function of six independent variables to that of two real functions of three independent variables, as a pure state corresponds to a complex function $\psi(x)$ or $\phi(p)$.

Since our df. f correctly gives expectation values of dynamical quantities at least of type (2.12b) as the mean values over f, the df. for pure state must satisfy the 'uncertainty relation' as the relation between the mean deviations of x and p:

$$\{\langle x_i^2\rangle_f-(\langle x_i\rangle_f)^2\}\{\langle p_i^2\rangle_f-(\langle p_i\rangle_f)^2\}\geq\varepsilon=\hbar^2/4. \tag{4.1}$$

This relation, being non-linear in f, is a necessary condition for the correspondence of f to a pure state, but not a sufficient one.

The condition that a mixed state should in particular fall into a pure state can be expressed, in terms of the density matrix, as $\rho^2=\rho$, i.e.,

$$\int\rho(x, x'')\rho(x'', x')dx''=\rho(x, x'). \tag{4.2}$$

Transforming this relation according to (2.6) we may immediately obtain the pure state condition for f, but then the result is not of a convenient form (see § 7)[*]; so we will prefer another way.

Now, if (4.2) is satisfied, there exists a suitable complex function $\psi(x)$ that makes $\rho(x, x')$ written as

$$\rho(x, x')=\psi(x)\psi^*(x'), \tag{4.3}$$

on account of (2.1), (2.2), and (2.3). It is because (4.2) with (2.3) requires $w_n=\delta_{nn_0}$ (with certain fixed n_0), when $\rho(x, x')$ is put in the form (2.4). But if (4.3) be satisfied, we have clearly

---

[*]   But there the condition is expressed in a form symmetrical in x and p.

$$\frac{\partial \rho}{\partial x_i} \frac{\partial \rho}{\partial x_k'} - \rho \frac{\partial^2 \rho}{\partial x_i \partial x_k'} = -\rho \frac{\partial \rho}{\partial x_k'} \frac{\partial}{\partial x_i} \log\left(\frac{\partial}{\partial x_k'} \log \rho\right) = 0 ; \tag{4.4}$$

$$(i, k = 1, 2, 3)$$

and conversely, if (4.4) be satisfied for every i and k, we get (4.3) by integrating (4.4) and by use of (2.1), (2.2). We can therefore *adopt the local relation* (4.4), *in place of the integral relation* (4.2), *as the pure state condition*, under the premises (2.1), (2.2), and (2.3).

Eq. (4.4) is of course compatible* with the equation of motion for ρ, which determines $\rho(x, x', t)$ from its initial value $\rho(x, x', t_0)$ uniquely, so it is sufficient to impose the pure state condition at a certain instant.

Now we can get the pure state condition for the ps. df. $f(x, p)$ satisfying the general conditions (2.8), (2.9) by transforming (4.4) into the relation as to f through (2.6). For the intermediary function $\bar{\rho}(x, y)$ of (3.3), (4.4) is written as

$$\left(\frac{1}{2}\frac{\partial}{\partial x_i} - \frac{\partial}{\partial y_i}\right)\bar{\rho} \cdot \left(\frac{1}{2}\frac{\partial}{\partial x_k} + \frac{\partial}{\partial y_k}\right)\bar{\rho}$$

$$= \bar{\rho} \cdot \left(\frac{1}{2}\frac{\partial}{\partial x_i} - \frac{\partial}{\partial y_i}\right)\left(\frac{1}{2}\frac{\partial}{\partial x_k} + \frac{\partial}{\partial y_k}\right)\bar{\rho}, \tag{4.5}$$

which, by the fourier transformation, *separates* into the following real and imaginary part equations :

$$(p_i f) * (p_k f) - f * (p_i p_k f) = -\varepsilon\left(\frac{\partial f}{\partial x_i} * \frac{\partial f}{\partial x_k} - f * \frac{\partial^2 f}{\partial x_i \partial x_k}\right), \tag{4.6}$$

$$\frac{\partial f}{\partial x_i} * (p_k f) - \frac{\partial f}{\partial x_k} * (p_i f) = f * \left(p_k \frac{\partial f}{\partial x_i} - p_i \frac{\partial f}{\partial x_k}\right), \tag{4.7}$$

where the notation such as $f_1 * f_2$ means the *convolution* with respect to p, i.e.,

$$f_1(x, p) * f_2(x, p) \equiv \int f_1(x, p') f_2(x, p - p') dp'$$

$$= \int f_1(x, p') f_2(x, p'') \delta(p - p' - p'') dp' dp''. \tag{4.8}$$

We have thus obtained the subsidiary conditions for the correspondence to a pure state, which have following properties :

(i)   The condition consists of six relations (4.6) symmetrical in i and k and three real anti-symmetrical ones (4.7). Since f is a scalar, (4.6) is a tensor equation and (4.7) a vector equation, both covariant to coordinate transformations.

(ii)   They are 'kinematical' relations independent of the dynamical characteristics of the system. Planck's constant which appeared in the time development equation (3.5) with (3.11) also enters the first condition (4.6), both in the form $\varepsilon \equiv \hbar^2/4$. Thus in our

---

*)   See the foot-note of appendix B.

formulation of quantum mechanics Planck's constant has to play such *twofold roles*. We may call (4·6) the 'quantum condition'** for the ps. en. formulation, and (4·7) the 'irrotationality condition' (see below).

(iii)   The conditions are surely not linear in f, but are again integral equations quadratic in f. If we have two distributions f_1 and f_2 satisfying both (4·6) and (4·7), the distribution obtained by their superposition, (which corresponds to mixing of states), no longer satisfies these conditions, as it should not.

We could adopt, in place of (4·4), the similar relation for $\rho(\boldsymbol{p}, \boldsymbol{p}')$:

$$\frac{\partial \rho}{\partial p_i} \frac{\partial \rho}{\partial p_k'} - \rho \frac{\partial^2 \rho}{\partial p_i \, \partial p_k'} = 0, \tag{4·9}$$

which, by fourier transformation using (2·7), leads to relations :

$$\left| \begin{array}{l} (x_i f) * (x_k f) - f * (x_i \, x_k f) = -\varepsilon \left(\dfrac{\partial f}{\partial p_i} * \dfrac{\partial f}{\partial p_k} - f * \dfrac{\partial^2 f}{\partial p_i \, \partial p_k} \right), \hfill (4·10) \\[3mm] \dfrac{\partial f}{\partial p_i} * (x_k f) - \dfrac{\partial f}{\partial p_k} * (x_i f) = f * \left(x_k \dfrac{\partial f}{\partial p_i} - x_i \dfrac{\partial f}{\partial p_k} \right), \hfill (4·11) \end{array} \right.$$

where the convolution is to be taken in respect to x in place of p. This set of relations is an alternative form of the pure state condition, being equivalent to the set (4·6) and (4·7).

(**b**)   Our next task is to re-express the conditions (4·6) and (4·7) in another forms. For that purpose, first, we integrate (4·6) and (4·7) throughout over the p-space, which procedure we shall call the '*projection* of the relations onto the coordinate space', and employ the factorization formula for the convolution :

$$\int (f_1 * f_2) \, d\boldsymbol{p} = \int f_1 d\boldsymbol{p} \int f_2 d\boldsymbol{p}. \tag{4·12}$$

Then we get from (4·6)

$$P_i P_k - P \cdot P_{ik} = \varepsilon (P \partial_i \partial_k P - \partial_i P \cdot \partial_k P), \qquad (\partial_i \equiv \partial/\partial x_i) \tag{4·13}$$

and from (4·7)

$$\partial_i P \cdot P_k - \partial_k P \cdot P_i = P(\partial_i P_k - \partial_k P_i), \tag{4·14}$$

where P is what was given by (2·10), and

$$\left| \begin{array}{l} P_i(\boldsymbol{x}) = \int p_i f(\boldsymbol{x}, \boldsymbol{p}) d\boldsymbol{p}, \hfill (4·15) \\[3mm] P_{ik}(\boldsymbol{x}) = \int p_i p_k f(\boldsymbol{x}, \boldsymbol{p}) d\boldsymbol{p} \hfill (4·16) \end{array} \right.$$

are distribution moments with respect to momentum components of first and second orders,

---

**) It is to be noted that the "quantum condition" in the old quantum theory determines stationary (pure) states, while our 'quantum condition' selects pure states out of mixtures.

respectively. Further, by introducing the mean values of p_i and $p_i p_k$ at each space point,

$$p_i(x) = P_i/P, \qquad p_i p_k(x) = P_{ik}/P, \tag{4.17}$$

(4.13) and (4.14) are simplified into

$$\overline{p_i\,p_k} - \bar{p}_i \cdot \bar{p}_k = -\varepsilon\; \partial^2\mathfrak{P}/\partial x_i \partial x_k, \qquad (\mathfrak{P} \equiv \log P) \tag{4.18}$$

$$\text{curl } \bar{\boldsymbol{p}} = 0. \tag{4.19}$$

Eq. (4.18) means the relation that the dispersion tensor of momentum, $\overline{p_i p_k} - \bar{p}_i \cdot \bar{p}_k$, at a point x should be connected with the space derivative of the space density at the point*, while (4.19) implies the irrotationality of the mean momentum field $\bar{\boldsymbol{p}}(x)$.

Next, we multiply (4.6) by p_j, and then again project the result onto the coordinate space to obtain

$$P_{ij}\,P_k + P_i\,P_{jk} - P_i P_{jl} - P \cdot P_{ijk}$$

$$= \varepsilon\{(P_j\,\partial_i\,\partial_k P + P\,\partial_i\,\partial_k\,P_j) - (\partial_i\,P_j\,\partial_k\,P + \partial_k\,P_j \cdot \partial_i\,P)\}. \tag{4.20}$$

By use of (4.18) and (4.19), this can be transformed into relations symmetrical** in $i, j,$ and k:

$$\overline{p_i p_j p_k} - \bar{p}_i \cdot \bar{p}_j \cdot \bar{p}_k = -\varepsilon\{(\textstyle\sum_{cyc} \bar{p}_i\,\partial_j\,\partial_k)\mathfrak{P} + \partial_i\,\partial_j\,\bar{p}_k\}, \tag{4.21}$$

which are ten relations for the (symmetrical) moment tensor of the third order:

$$P_{ijk}(x) = P \cdot \overline{p_i\,p_j\,p_k} = \int p_i\,p_j\,p_k f(x,\,\boldsymbol{p})\,d\boldsymbol{p}. \tag{4.22}$$

Continuing similar procedures on (4.18), we get in succession relations for successively higher moments. For instance, for the forth order moments we get 18 symmetrical relations:

$$\overline{p_i p_j p_k p_l} - \bar{p}_i \bar{p}_j \bar{p}_k \bar{p}_l = -\varepsilon(\textstyle\sum \bar{p}_i\,\partial_j\,\partial_k\,\bar{p}_l + \sum \bar{p}_i\,\bar{p}_j\,\partial_k\,\partial_l\mathfrak{P})$$

$$+ \varepsilon^2(\textstyle\sum_{cyc} \partial_i\,\partial_j\,\mathfrak{P}\cdot\partial_k\,\partial_l\,\mathfrak{P} + \partial_i\,\partial_j\,\partial_k\,\partial_l\,\mathfrak{P}). \tag{4.23***}$$

On the other hand similar operations on (4.7) lead to no new relations. We can now take, as the pure state condition, (4.6) and (4.19) in place of (4.6) and (4.7), or take (4.19) and all of the relations for successively higher moments: (4.18), (4.21) (4.23),···, which reduce every moment in and above second order to even order space derivatives of $\mathfrak{P}$ and $\bar{p}$ (i.e., the zeroth and first order moments). Thus the pure state conditions imply such restrictions to the df. that leave the zeroth moment free, restrict the first moments to being irrotational, and then uniquely determine the higher moments in terms of the zeroth and first moments. Thus we find that the functional

---

*) It is to be noted that the mean deviation of momentum components $\overline{p_i^2} - (\bar{p}_i)^2$ at a space point is not necessarily positive for our ensemble whose df. $f(x,\,p)$ is not necessarily positive, and is equal to $-\varepsilon\partial^2\mathfrak{P}/\partial x_i^2$ for pure states.

**) $\partial_i\partial_j f_k$ is symmetric by virtue of (4.19).

***) In the right side of (4.23) the first summation contains 4 terms, the second summation 6 terms, and the third summation 3 terms.

space of the df. $f(x, p)$ is indeed just limited to that of two real functions of x by means of our pure state conditions.

In a similar fashion we may also transform the conditions (4·10) and (4·11), by use of the projection onto the momentum space. We then obtain the relations in which the roles of x and p in (4·18), (4·19), (4·21),··· are just exchanged. For instance, in place of (4·18) and (4·19), we get

$$\begin{cases} \overline{x_i x_k} - \bar{x}_i \cdot \bar{x}_k = -\varepsilon\ \partial^2 \mathfrak{Q}/\partial p_i\, \partial p_k, & (\mathfrak{Q} \equiv \log Q) & (4·24) \\ \mathrm{curl}_p \bar{x} = 0, & & (4·25) \end{cases}$$

where

$$\begin{cases} \bar{x}_i(p) = Q_i/Q, & Q_i(p) = \int x_i f\, dx, \\ \overline{x_i x_k}(p) = Q_{ik}/Q, & Q_{ik}(p) = \int x_i x_k f\, dx. \end{cases} \tag{4·26}$$

Finally there is a problem more general than that of the pure state condition, i.e., to give a measure to 'the degree of mixture' for an arbitrary distribution $f(x, p)$. It may be achieved by the introduction of entropy $\sum$ defined by

$$\sum = \mathrm{Sp}\,(\rho \log \rho) = \sum_{n=1}^{\infty} (-1)^n/n \cdot \mathrm{Sp}\,\{(\rho^2 - \rho)\,(\rho - 1)^{n-1}\}\,.$$

In terms of f, however, $\sum$ should take a too complicated form.

(c) In our formulation quantum-mechanical change of pure state is described by the quantum Liouville equation (3·5) for the df. $f(x, p, t)$ and the subsidiary conditions (4·6) and (4·19) which is compatible with (3.5). Naturally this description is equivalent to the usual one in terms of the wave function $\psi(x, t)$ that obeys the Schrödinger equation : First, if the latter is given, we construct the df. $f(x, p\ t)$ by use of $\rho(x, x', t) = \psi(x, t)\psi(x', t)^*$ and (2.6), then this df. clearly satisfies the quantum Liouville equation and the pure state conditions. Conversely, if a df. $f(x, p, t)$ satisfying the quantum Liouville equation and the pure state condition is first given, we produce the corresponding $\rho(x, x', t)$ by use of (3.3). Then this ρ must be hermitian and satisfy (4.4), and hence factorize as (4.3) i.e.,

$$\rho(x, x', t) = \psi(x, t)\psi^*(x', t). \tag{4·27}$$

This determines the wave function ψ up to an arbitrary phase depending upon time only ; that is $\psi(x, t)$ can be written as

$$\psi(x, t) = \psi_0(x, t)e^{i\Lambda(t)}, \tag{4·28}$$

where $\Lambda(t)$ is an arbitrary function. Furthermore, $\psi_0(x, t)$ here can be taken to be the Schrödinger equation :-

$$K[\psi] \equiv (\hbar/i \cdot \partial/\partial t - \hbar^2/2m \cdot \varDelta + V(x))\psi = 0, \tag{4·29}$$

because the above ρ must satisfy its equation of motion (3.2) and so $\psi(x, t)$ in (4.27) satisfies

$$K[\psi(x)]\psi^*(x') - K^*[\psi^*(x')]\psi(x) = 0.$$

Consequently we can uniquely determine the wave function $\psi(x, t)$ that corresponds to the given df. from (4.27) and the '*supplementary condition*' (4.29).

§ 5.   Equivalence between the formulation in terms of phase-space ensemble and that in terms of configuration space ensemble

(a)   We have established that a quantum-mechanical change of pure state is described by the ps. df. obeying quantum Liouville equation and certain subsidiary conditions, as well as by the wave function obeying the Schrödinger equation.   On the other hand quantum-mechanical motion can also be represented by certain trajectory ensemble in configuration space (cs. en.), as was analysed in our previous papers[4].   Therefore, the representation of quantum-mechanical motion for the case of pure state in terms of ps. en. must be equivalent to that in terms of the cs. en.   We shall now examine this point in a direct manner, taking out the correspondence between both formulations.

First we briefly recapitulate the method of the cs. en. for the case of a single particle under consideration.   This method represents a quantum-mechanical state of wave function

$$\psi = R e^{iS/\hbar} \qquad (R, S : \text{real}) \tag{5.1}$$

with an ensemble which consists of a probability distribution of a particle in the density

$$P(x) = R(x)^2, \tag{5.2}$$

the particle momentum p being uniquely correlated with its position x by

$$p(x) = \nabla S(x). \tag{5.3}$$

Thus the momentum field satisfies

$$\text{curl } p = 0 ; \tag{5.4}$$

and the ensemble has a particular *phase-space* distribution :

$$f_c(x, p) = R(x)^2 \cdot \delta(p - \nabla S(x)). \tag{5.5}$$

The Schrödinger equation is written, in terms of R and S, as

$$\partial S/\partial t + 1/2m \cdot (\nabla S)^2 + V - (\hbar^2/2m)\Delta R/R = 0, \tag{5.6}$$

$$\partial(R^2)/\partial t + \text{div}(R^2 \nabla S/m) = 0, \tag{5.7}$$

of which the latter gives the equation of continuity for the cs. en. :

$$\partial P/\partial t + \text{div}(Pv) = 0, \qquad (v = p/m) \tag{5.8}$$

while the former leads to the equation of motion for a particle of the ensemble :

$$m \, dv/dt = -\nabla(V + V'), \tag{5.9}$$

with

$$V' = -\hbar^2/2m \cdot \Delta R/R. \tag{5.10}$$

Thus the temporal development of the ensemble can be regarded as is built up through the process in which each point of the ensemble moves along a continuous path with the momentum $(5\cdot3)$ at each instant, accelerated not only by the external potential V but also by the additional 'quantum potential' $(5\cdot10)$. In place of $(5\cdot9)$ we may also adopt the relation of momentum conservation:

$$\frac{\partial(Pp_i)}{\partial t}+\frac{1}{m}\sum_k\frac{\partial(Pp_ip_k)}{\partial x_k}=-P\frac{\partial V}{\partial x_i}+\sum_k\frac{\partial\sigma_{ik}}{\partial x_k}, \qquad (5\cdot11)$$

where σ_{ik} means the 'quantum stress',

$$\sigma_{ik}=\hbar^2/4m\cdot P\ \partial^2(\log P)/\partial x_i\,\partial x_k. \qquad (5\cdot12)$$

We could thus associate an ensemble of trajectories satisfying

$$\left.\begin{array}{l}\text{equation of motion }(5.9),\ (\text{or }(5.11)),\\[4pt]\text{equation of continuity }(5.8),\\[4pt]\text{subsidiary condition }(5.4),\end{array}\right\} \qquad (A)$$

with a wave function $\psi(x,\ t)$ satisfying the Schrödinger equation, by means of (5.2) and (5.3). Conversely, any ensemble of trajectories that satisfies (A) corresponds to a quantum-mechanical change of state as follows: Given a solution of (A), $P(x,t)$ and $p(x,t)$, we can determine R and ∇S by (5.2) and (5.3), and so $S(x,t)$ itself can also be determined up to an arbitrary function of time, $\Lambda(t)$. But this arbitrariness is excluded by imposing on S the 'supplementary condition' that S should satisfy (5.6), which condition is clearly compatible with the equation of motion (5.9). Thus we uniquely get the wave function $\psi=Re^{iS/\hbar}$ satisfying the Schrödinger equation*.

(**b**) We shall now explicitly show the correspondence between our ps. en. representation and the cs. en. representation outlined just now in two steps.

i) In the first place, given a phase space distribution $f(x,p,t)$ corresponding to a change of pure state, the corresponding cs. en. can be produced by 'projecting' the ps. en. onto the coordinate space. This means that we introduce the cs. en. which consists of the density and momentum fields, $P(x,t)$ and $p(x,t)$, derived from the ps. df. $f(x,p,t)$, by (2.10) and (4.17) with (4.15); in other words we eliminate the momentum dispersion at each space point x in the ps. en., adopting the average momentum and the total density at each point x. We can then show that the cs. en. obtained satisfies the condition (A) and represents the same quantum-mechanical change of state.

To show this, first we project the quantum Liouville equation (3.5) for f onto the coordinate space to obtain

$$\partial P/\partial t+\mathrm{div}\,(P\bar{p}/m)=0, \qquad (5\cdot13)**$$

---

*) Apart from a physically meaningless arbitrary additive constant in the phase.

**) The stochastic transitions in momentum have no effect on the time change of the total space density, since they occur with positive as well as negative probabilities for various jumps and cancell out in the sum.

which expresses the continuity equation (5.8) for the cs. en. $(P, \ \bar{\boldsymbol{p}})$ derived by projection.

Next we project the momentum conservation relation for the ps. en.,

$$\frac{\partial(p_i f)}{\partial t} + \frac{1}{m}\sum_k p_i p_k \frac{\partial f}{\partial x_k} = \int p_i f(\boldsymbol{x}, \boldsymbol{p} - \boldsymbol{p}') f(\boldsymbol{x}, \boldsymbol{p}') d\boldsymbol{p}', \tag{5.14}$$

obtainable from (3.5), and get

$$\frac{\partial(P\bar{p}_i)}{\partial t} + \frac{1}{m}\sum_k \frac{\partial}{\partial x_k}(P \ \overline{p_i p_k}) = -P\frac{\partial V}{\partial x_i}. \tag{5.15*)}$$

Here we take into account one of the pure state conditions (4.18) for the original phase space distribution, then the second term in (5.15) may be rewritten as

$$\frac{1}{m}\sum_k \frac{\partial}{\partial x_k}(P \ \overline{p_i p_k}) = \frac{1}{m}\sum_k \frac{\partial}{\partial x_k}(P \ \bar{p}_i \cdot \bar{p}_k) - \frac{\hbar^2}{4m}\sum_k \frac{\partial}{\partial x_k}\left(P\frac{\partial^2 \log P}{\partial x_i \ \partial x_k}\right). \tag{5.16}$$

This means that the contribution from the convection term in the ps. en. picture transforms to the convection term in the derived cs. en. picture plus the extra momentum flow such as is ascribable to the occurrence of the 'quantum stress' (5.12). Now (5.15) with (5.16) is exactly the momentum conservation (5.11) for the derived cs. en. $(P, \ \bar{\boldsymbol{p}})$, and so we can also obtain the equation of motion (5.9).

Furthermore another one of pure state conditions, (4.19), for the ps. en. immediately warrants the subsidiary condition (5.4) for the derived cs. en. $(P, \ \bar{\boldsymbol{p}})$. Thus we can conclude that the cs. en. produced from the original ps. en. by projection is in fact a *possible* one satisfying (A).

ii)   Conversely, if we are given a cs. en. specified with $P(\boldsymbol{x})$ and $\bar{\boldsymbol{p}}(\boldsymbol{x})$, satisfying (A), we can consider many a phase space distribution $f(\boldsymbol{x}, \boldsymbol{p})$ which can yield that cs. en. $(P, \ \bar{\boldsymbol{p}})$ by projection. However, from P and $\bar{\boldsymbol{p}}$ we determine successively quantities, $\overline{p_i p_k}, \ \overline{p_i p_j p_k}, \cdots$, according to (4.18), (4.21),$\cdots$, and then we can determine a ps. df. $f(\boldsymbol{x}, \boldsymbol{p})$ *uniquely*, such that it takes those values, $P, \ \bar{p}_i, \ \overline{p_i p_k}, \ \overline{p_i p_j p_k}, \ \cdots$, as its successively higher moments, since now every order moment is specified for the df. This ps. df.** is the only one that yields $(P, \ \bar{\boldsymbol{p}})$ by projection and satisfies at the same time the pure state conditions. We can further show that this ps. df. fulfills the quantum Liouville equation.

') In this equation for the time change of mean momentum, the effects of the stochastic transitions in momentum induced by the external potential V are reduced to the classical value $-P\partial V/\partial x_i$.

**) The explicit form of this df. can be written down as

$$f(\boldsymbol{x}, \boldsymbol{p}, t) = \frac{1}{(2\pi\hbar)^3}\int\left[P\left(\boldsymbol{x} - \frac{\boldsymbol{y}}{2}, \ t\right)P\left(\boldsymbol{x} + \frac{\boldsymbol{y}}{2}, \ t\right)\right]^{1/2}\exp\frac{i}{\hbar}\left[S\left(\boldsymbol{x} - \frac{\boldsymbol{y}}{2}, \ t\right) - S\left(\boldsymbol{x} + \frac{\boldsymbol{y}}{2}, t\right) + \boldsymbol{p}\boldsymbol{y}\right]d\boldsymbol{y},$$

where S is a potential function for $\bar{\boldsymbol{p}}$, i.e., $\nabla S = \bar{\boldsymbol{p}}$.

In all above the equivalent correspondence between ps. en. and cs. en. has perfectly been verified.

(**c**)   In a previous paper[5], we inquired whether the quantum potential (5.10) in the cs. en. formulation could be analysed into a mechanism like any Markoff process underlying. The problem is now explained more clearly : From the viewpoint of the ps. en. formulation, the quantum potential can be regarded as an *apparent force* appearing as the result of projecting on to the coordinate space the ps. en. that satisfies the pure state condition and changes according to a sort of Markoff process.   We cannot, however, regard the latter picture as a literally real one any more than the picture of trajectory ensemble under quantum potential, on account of the inevitable appearance of negative probabilities.   Furthermore it is also to be noted that the ps. en. could not yield mean values correctly for certain quantities, and that the pure state condition was in a sense ad hoc.

Notwithstanding, our formulation of quantum mechanics in terms of Markoff-like picture might further tempt the idea of some hidden mechanism of irregular external disturbances which vanish in the average yet make the particle momentum fluctuate, acting on particle irrespective of its momentum [see (ii) of § 3 (a)].   But the features of stochastic transitions stated in § 3 (a) do not allow to construct any such model *realistically*.

Recently Weizel[14] attempted to derive the quantum potential from certain stochastic process based on some model.   He proceeded in a considerably different fashion, but the nature of his method may also be illuminated from our viewpoint which may be more far-reaching than his method, standing upon the systematic formulation of quantum mechanics in terms of the ps. en.

By the way, the mean kinetic energy of particle for the observer moving with the mean velocity in the ps. en. is $E_1 = 1/2m \cdot \{\overline{p^2} - (\overline{p})^2\}$, which is not necessarily positive, and becomes $E_1 = -\varepsilon/2m \cdot \Delta\mathfrak{P}$ for a pure state due to (4·18).   On the other hand the mean pressure p_r in the corresponding cs. en. is[4] $p_r = -\frac{1}{3}\sum_i \sigma_{ii} = -\frac{1}{3}\frac{\varepsilon}{m}P\Delta\mathfrak{P}$, and so we have the relation $p_r = 2/3 \cdot PE_1$.

This shows that the pressure in the cs. en. results from the momentum dispersion of the underlying ps. en. just in the same manner as the pressure of ideal gas results from the thermal motion of molecules.

(**d**)   We have explained that a cs. en. can be looked upon as the projection of a ps. en. for pure state.   It is to be remarked that as a result of such contraction the cs. en. gets free from an unrealistic property of the ps. en., i.e., negative probabilities, but at the same time it partly loses the property of the ps. en. to give correctly the quantum-mechanical expectation values as the ensemble averages for most of usual quantities :   The cs. en. defined with density and momentum fields, $P(x)$ and $\nabla S(x)$, yields the mean values for a quantity $A(x, p)$ as

$$\langle A \rangle_o = \int A(x, \nabla S) P(x) \, dx,$$

which agrees with $(2 \cdot 12a)$ for quantities of zeroth or first order in p but not for those of second or higher orders[1].

As for the ps. en., relations known to hold between quantum mechanical expectation values of some physical quantities usually hold also with the understanding that the average over the ps. en. is to be taken, since it yields the expectation values correctly as the ensemble averages for most of usual physical quantities. An example is the 'uncertainty relation' stated in § 4 (a); another one is the 'Ehrenfest's theorem': Integrating $(5 \cdot 15)$ throughout over the x-space we get

$$a^2 \langle x_i \rangle_f / dt^2 = d \langle p_i \rangle_f / dt = -\langle \partial V / \partial x_i \rangle_f,$$

which is valid for general mixed states. In the particular case of pure states, this relation may also be written, in terms of the quantities in the cs. en., as

$$m \frac{d^2}{dt^2} \int x P \, dx = \frac{d}{dt} \int P \bar{p} \, dx = -\int P \nabla V \, dx.$$

Still another example is the 'virial theorem',

$$d \langle xp \rangle_f / dt = 1/m \cdot \langle p^2 \rangle_f + \langle x \nabla V \rangle_f,$$

and also the variational theorem, into which we do not enter here.

§ 6. Applications of the ps. en. formulation to some elementary examples of pure state cases

(a) Distributions without correlation and the diffusion of wave packets

i) We shall ask whether there can exist any pure state distributions having no correlation at all between particle position and momentum. Such distribution must factorize as

$$f(x, p) = P(x) Q(p), \tag{6.1}$$

and therefore satisfies one of the pure state conditions $(4 \cdot 19)$ from the outset. Since it must further satisfy $(4 \cdot 6)$, we obtain

$$\left\{ \begin{array}{ll} \dfrac{\partial^2 \log P}{\partial x_i \, \partial x_k} = -\dfrac{1}{2} \alpha_{ik}, & (\alpha_{ik} = \alpha_{ki} = \text{real const.}) \tag{6.2} \\[2ex] (p_i Q) * (p_k Q) - Q * (p_i p_k Q) = -(1/2) \varepsilon \alpha_{ik} Q * Q. \tag{6.3} \end{array} \right.$$

Eq. $(6 \cdot 2)$ yields

$$P = \text{const. } \exp\left[-\sum_{i, k} \alpha_{ik} x_i x_k + \sum_i \gamma_i x_i \right], \tag{6.4}$$

which can be led, by a suitable orthogonal coordinate transformation,

$$x_i' = \sum_k c_{ik} \, x_k, \tag{6.5}$$

to the form

$$P(x) = \text{const. } \exp\left[-\sum_i \alpha_i (x_i' - a_i')^2\right], \tag{6.6}$$

where each α_i should be positive so that P should be normalizable.

Next[*], re-expressing (6.3) in terms of the fourier transform for $Q(p)$,

$$\chi(\eta) = \int Q(p) e^{ip\eta} \, dp, \tag{6.7}$$

we get

$$-\frac{\partial\chi}{\partial\eta_i} \frac{\partial\chi}{\partial\eta_k} + \chi \frac{\partial^2\chi}{\partial\eta_i\partial\eta_k} = -\frac{\varepsilon}{2}\alpha_{ik}\,\chi^2. \tag{6.8}$$

This is a equation similar to (6.2), and can be integrated into

$$\log\chi = -\varepsilon\sum_i\alpha_i(\eta_i' - \delta_i^\dagger)^2 + \text{const.},$$

where η_i's are the new components produced by the coordinate transformation (6.5) applied to η_i's. We have now

$$Q(p) = \text{const. } \exp\left[-\sum_i (1/4\varepsilon\alpha_i)(p_i' - b_i')^2\right], \tag{6.9}$$

where $p_i' = \sum_k c_{ik}\,p_k$. Altogether, our distribution must be of the form

$$\left\{\begin{array}{ll} f(x, p) = \text{const. } \exp\left[-\sum_i\{\alpha_i(x_i' - a_i')^2 + \beta_i(p_i' - b_i')^2\}\right], & (6.10a) \\ \alpha_i\beta_i = 1/\hbar^2, & (6.10b) \end{array}\right.$$

namely it is Gaussian for each freedom of position and momentum components in a suitable coordinate system, with the relation (6.10b). It has the minimum 'uncertainty product':

$$\langle(x_i - \langle x_i\rangle_f)^2\rangle_f \cdot \langle(p_i - \langle p_i\rangle_f)^2\rangle_f = \varepsilon.$$

The wave function corresponding to (6.10) can be obtained according to the procedure of § 4 (c): First we derive

$$\bar{\rho}(x, y) = \int f(x, p) e^{-ipy/\hbar} \, dp = P(x) \cdot \chi(-y/\hbar)$$
$$= \text{const. } \exp\left[-\sum_i\{\alpha_i(x_i' - a_i')^2 + (1/4) \, \alpha_i y_i'^2 + (i/\hbar) \, b_i' y_i'\}\right],$$

from which the density matrix is obtained as

$$\rho(x, x_1) = \text{const. } \exp\sum_i[-\alpha_i/2 \cdot \{(x_i' - a_i')^2 + (x_{1i}' - a_i')^2\} + (i/\hbar)(b_i'(x_i' - x_{1i}'))],$$

so the wave function takes the form

$$\psi(x) = \text{const. } \exp\left[-\sum_i\frac{\alpha_i}{2}(x_i' - a_i')^2 + i(\sum_i b_i' x_i'/\hbar + \Lambda(t))\right], \tag{6.11}$$

---

[*] From here on it would be simpler to proceed as follows: The other form of the pure state condition, (4.10), yields $\partial^2 \log Q/\partial p_i \, \partial p_k = -\beta_{ik}/2$ which is similar to (6.2), and the condition (4.18) further requires (6.10b).

which is nothing but the general form of the 'minimum wave packet'.

ii)   Next we shall ask whether or not the form of the distribution (6.10) (i.e., the property of having no correlation) be conserved during the course of time.   First, taking up the case of free particle*, we suppose that a distribution of type (6.10),

$$f_0(x, p) = \frac{1}{\pi\hbar} \exp\left[-\alpha(x-a)^2 - \frac{1}{\alpha\hbar^2}(p-b)^2\right],\qquad (6 \cdot 12)$$

(here considering one-dimensional case for simplicity) occurs at a time $t=t_0$. In this case, as was stated in § 3 (a), each point in the ps. en. moves classically with its *respective* constant velocity, and the distribution at time t is given by

$$f(x, p, t) = f_0(x - p/m \cdot (t-t_0), p)$$

$$= \frac{1}{\pi\hbar} \exp\left[-\alpha\left(x - a - \frac{p}{m}(t-t_0)\right)^2 - \frac{1}{\alpha\hbar^2}(p-b)^2\right].\qquad (6 \cdot 13)$$

But this is no longer of type (6.10), and indicates that the correlation grows with time. The phase space distribution (6.13) gives at once the space density in the well-known form

$$P(x) = \int f dp = \text{const.}\left[\frac{a}{1+\{(\alpha\hbar/m) \cdot (t-t_0)\}^2}\right]^{1/2} \exp\left[-\frac{a(x-a)^2}{\{1+(\alpha\hbar/m) \cdot (t-t_0)\}^2}\right],$$

showing that the distribution in coordinate space diffuses with the passage of time. Thus the so-called "diffusion of wave-packet" in this case is, from our viewpoint, simply a result of the fact that each point of the ps. en. performs the purely classical motion with its *respective* momentum[15].   In other words the simple circumstance that the particles which move faster cover the greater distances brings about the spreading of the packet and at the same time introduces correlation.

(b)   Linear oscillator

i)   The hamiltonian is

$$H(x, p) = 1/2 \cdot (p^2/m + m\omega^2 x^2),$$

and the quantum Liouville equation is identical with the classical one:

$$\partial f/\partial t + p/m \cdot \partial f/\partial x - m\omega^2 x \, \partial f/\partial p = 0,\qquad (6 \cdot 14)$$

of which the general solution is

$$f = F(p^2 + (m\omega x)^2, \ -p \sin \omega t + m\omega x \cos \omega t).\qquad (6 \cdot 15)$$

Now in case of oscillator, if there can exist a pure-state distribution having no correlation during the course of time, such distribution must be of the form (6.12) and at the same time consistent with the form (6.15).   This determines the parameters in (6.12) such that

---

*)   For the case of oscillator, see (b) i).

$$\alpha = m\omega/\hbar, \quad a = a_0 \cos \omega t, \quad b = -m\omega a_0 \sin \omega t, \quad (a_0: \quad \text{const.}) \tag{6.16}$$

and so we must have the df. of the form,

$$f(x, p, t) = \frac{1}{\pi\hbar} \exp\left[-\frac{m\omega}{\hbar}(x - a_0 \cos \omega t)^2 - \frac{1}{m\omega\hbar}(p + a_0 m\omega \sin \omega t)^2 \right]. \tag{6.17}$$

This distribution takes its maximum value at the point $(x = a_0 \cos \omega t, \ p = -a_0 m\omega \sin \omega t)$ in phase space, decreasing around it in Gaussian manner. With the passage of time the distribution rotates, *keeping its form rigidly*, along an ellipse around the origin of phase space with the angular frequency ω.

The corresponding wave function[16] is obtained by inserting (6.16) into (6.11) as

$$\psi(x, t) = \exp\left[-\frac{m\omega}{2\hbar}(x - a_0 \cos \omega t)^2 - \frac{i}{\hbar} m\omega x a_0 \sin \omega t + iA(t) \right]. \tag{6.18}$$

Its phase, $S = -m\omega/\hbar \cdot x a_0 \sin \omega t + A(t)$, involves a yet undetermined part $A(t)$, which is determined, by the 'supplementary condition' (5.6) for S, as

$$A(t) = \frac{1}{4} \frac{m\omega}{\hbar} a_0^2 \sin 2\omega t - \frac{1}{2}\omega t. \tag{6.19}$$

ii) As previously stated, for quadratic potential such as in case of oscillator, the ps. en. develops classically. On the other hand, in such potential, also in the cs. en. formulation, we have particular solutions[1] which consist of purely classical trajectories, with quantum potential vanishing, and are determined by the following equations:

$$\begin{cases} \dot{S} + 1/2m \cdot (\nabla S)^2 + V = 0, & (6.20) \\ \nabla(\Delta S) = 0, & (6.21) \\ P(t) = \text{const} \cdot \exp\left[-(1/m)\int^t \Delta S \, dt \right]. & (6.22) \end{cases}$$

Now, such a cs. en. corresponds to a ps. en. with the df.

$$f(x, p, t) = \frac{1}{(2\pi\hbar)^3} P(t) \int \exp \frac{i}{\hbar}\left[S\left(x - \frac{y}{2}, t\right) - S\left(x + \frac{y}{2}, t\right) + py \right] dy, \tag{6.23}$$

which can be simplified in one-dimensional case as

$$f(x, p, t) = P(t) \cdot \delta(p - \partial S/\partial x), \tag{6.24}$$

because from (6.21) we have $S(x - y/2) - S(x + y/2) = -y \, \partial S/\partial x$. The expression (6.24) (together with (5.5)) shows that for such particular quantum-mechanical motion the ps. en. and the cs. en. become identical with each other.

In the case of oscillator, a solution satisfying (6.20), (6.21), and (6.22) is given by[4]

$$\begin{cases} P = \text{const}/\sin\omega\tau, & (\tau = t - t_0) & (6.25a) \\ S = (m\omega/2\sin\omega\tau)\{(x^2 + x_0^2)\cos\omega\tau - 2xx_0\}, & (6.25b) \end{cases}$$

and so by (6.24) the corresponding df. becomes

$$f(x, p) = f_c(x, p) = \text{const. } \delta(p \sin \omega\tau - m\omega(x \cos \omega\tau - x_0)). \quad (6 \cdot 26)^*$$

This expresses an ensemble of trajectories flowing out x_0 at t_0 with *any* velocity. The corresponding wave function $P^{1/2} e^{iS/\hbar}$ is nothing but the propagation kernel $K(xt|x_0t_0)$. It is to be noted that the 'transition probability' of § 3 (b) is a different thing from the above df., and is given by (3.35) in this case, implying the distribution at t when a particle starts from x_0 at t_0 with a *definite* momentum p_0.

iii) We call a time-independent solution of (3.5) a *stationary distribution*, where the deviation of the distribution to be caused by the momentum transition with the transition probability $J(x, p'-p)$ is exactly compensated with the effect of the translation of particles.

A quantum-mechanical *energy eigen-state* corresponds to a *stationary distribution such as satisfies the pure state conditions* in our phase space formulation. It is, however, to be remarked that such distribution is *not suited* to be called "energy eigenstate" in our picture. Generally quantum-mechanical eigenstates for H or l_i cannot be specified as the ps. ensembles satisfying pure state conditions and having no dispersion in respect to respective quantities (cf. § 2). We shall next see this point for the case of oscillator.

As is well-known the stationary state wave functions for the oscillator are

$$\psi_n(x,t) = \left(\frac{\kappa}{2^n \pi^{1/2} n!}\right)^{1/2} e^{-\kappa^2 x^2/2} H_n(\kappa x) e^{-i(n+1/2)\omega t}, \qquad \kappa = (m\omega/\hbar)^{1/2} \quad (6 \cdot 27)$$

from which we can obtain the corresponding ps. dfs. as

$$f_n(x, p, t) = \frac{(-1)^n}{\pi\hbar} e^{-2H(x, p)/\hbar\omega} L_n(4H(x, p)/\hbar\omega). \quad (6 \cdot 28)^{**}$$

Each of these distributions has a constant density on an energy surface, as it should for a stationary distribution in case of any quadratic potential, and takes negative as well as positive values, giving the quantum-mechanical expectation value of energy correctly as

$$\langle H \rangle_{f_n} = \int H(x, p) f_n(x, p) dx \, dp = (n+1/2)\hbar\omega = E_n = \langle H \rangle_n.$$

However, it cannot give the expectation value correctly for any power of energy, $H^\nu (\nu \geq 2)$, because f_n distributes over an *area* in phase space.

The ps. distribution that yields the probability distribution correctly for H is clearly the

---

*) In this case the wave function and therefore the ps. df. are not normalizable; the const. in (6.22) and (6.26) are in reality infinitesimal.

**) $L_n(\zeta)$ is a Laguerre polynomial defined by $L_n(\zeta) = \sum_{k=0}^{n} \frac{(-1)^n}{k!} \binom{n}{k} \zeta^k$. These distribution functions $(-1)^n \pi\hbar f_n(x, p)$ $(n=0, 1, 2 \cdots)$ constitute a complete orthonormal set as functions of $(4/\hbar\omega) H(x, p)$ in the domain $(0, \infty)$.

one which concentrates on the ellipse, $H(x, p) = (n+1/2)\hbar\omega$, in phase space, i.e.,

$$f_n'(x, p) \propto \delta(H(x, p) - (n+1/2)\hbar\omega).$$

Such distribution cannot, however, correspond to any state, as is understood, from the fact that it does not satisfy the Wigner's condition $(7 \cdot 1)$.

§ 7. Remarks on the Wigner's condition and alternative forms of the pure state condition

(a) E.P. Wigner pointed out* that following two important conditions are further *necessary*, besides the conditions $(2 \cdot 9)$ and $(2 \cdot 19)$, for a ps. df. $f(x, p)$ to be a permissible one corresponding generally to a mixed state. They are

$$|f(x, p)| \leq (2/h)^3, \qquad (7 \cdot 1)$$

$$h^3 \int f(x, p)^2 dx \, dp \leq 1. \qquad (7 \cdot 2)$$

The condition $(7 \cdot 1)$ which is derivable by applying Schwarz inequality to $(2 \cdot 6)$ with $(2 \cdot 4)$ indicates that a possible df. should extend at least over a phase volume $(h/2)^3$, expressing in a certain degree the uncertainty principle for the general case of a mixed state. The left side of $(7 \cdot 2)$, being

$$h^3 \int f(x, p)^2 dx \, dp = \text{Sp } \rho^2 = \sum_n w_n^2 \qquad (7 \cdot 3)$$

(where w_n is the quantity defined by $(2 \cdot 4)$), becomes unity for a pure state and is smaller for a general mixed state. This quantity may be taken as giving a measure to the degree of mixing, though it is different from entropy defined in § 4 (b).

The Wigner's conditions (7.1) and (7.2) must be regarded, together with $(2 \cdot 19)$ as a part of the 'positivity condition' stated in § 2, which means the condition to be imposed upon any possible ps. df. corresponding to $(2 \cdot 3)$. These conditions (2.19), (7.1), and (7.2) may not yet constitute the sufficient one, since the positivity condition means that every w_n be non-negative and so it would require infinitely many inequalities of a type such as (7.2). Indeed, the condition (7.2), which is expressible in terms of density matrix as

$$\int \rho(x \, x') \rho(x' \, x) \, dx \, dx' \leq 1, \qquad (7 \cdot 2')$$

is merely a part of the condition

$$\int \rho(xx') \rho(x'x) dx' \leq \rho(x \, x) = P(x). \qquad (7 \cdot 4)$$

The latter condition $(7 \cdot 4)$ is also a *necessary* one, and is expressed in the language of the ps. df. as

$$2^3 \int f(x' \, p) f(x' \, p') \cos[2/\hbar \cdot (p'-p)(x'-x)] dp \, dp' \, dx'$$

$$\leq \int f(xp)\,dp = P(x). \qquad (7\cdot5)$$

The left side of $(7\cdot5)$ is non-negative, and therefore (7.5) means a severer condition than the first relation of (2.19), and at the same time involves $(7\cdot2)$. Similarly the second relation of (2.19) can better be replaced by

$$2^3 \int f(x\,p')f(x'\,p')\cos[2/\hbar\cdot(p'-p)(x'-x)]\,dx\,dx'\,dp' \leq Q(p). \qquad (7\cdot6)$$

(b)   Next we shall reconsider the pure state condition. Taking the equality sign in $(7\cdot2)$ we obtain

$$\hbar^3 \int f(x,p)^2\,dx\,dp = 1 \qquad (7\cdot2^\circ)$$

as a necessary condition for a ps. df. to correspond to a pure state. This condition has a different form from those obtained in § 4. Now we can show that the pure state condition can also be expressed as a series of infinitely many relations each of which is of a type similar to $(7\cdot2^\circ)$, the series including (7.2°) itself as its first relation. For that purpose we start from the original form of the pure state condition for density matrix, $(4\cdot2)$, in place of (4.4). Re-expressing (4.2) in terms of ps. df., we obtain

$$2^3 \int f\left(x'-\frac{y}{2},p\right)f\left(x'+\frac{y}{2},p'\right)\exp\frac{i}{\hbar}[2(p'-p)(x'-x)-(p+p')y]dx'\,dp\,dp'$$

$$= \int f(x,p)\exp(-2\,i/\hbar\cdot py)dp. \qquad (7\cdot7)$$

This complicated relation involving two continuous parameters x and y can be split into a series of infinitely many simpler relations by the following two-step procedure. First, integrating (7.7) over the whole x-space, we obtain

$$\hbar^3 \int f\left(x-\frac{y}{2},p\right)f\left(x+\frac{y}{2},p\right)\exp\left(-2\frac{i}{\hbar}py\right)dx\,dp$$

$$= \int f(x,p)\exp\left(-2\frac{i}{\hbar}py\right)dx\,dp. \qquad (7\cdot8)$$

Next we integrate $(7\cdot7)$ after multiplying x_i on both sides to obtain

$$\hbar^3 \int f\left(x-\frac{y}{2},p\right)\left\{\frac{\hbar}{2i}\frac{\partial}{\partial p_i}f\left(x+\frac{y}{2},p\right)+\left(x_i-\frac{y_i}{2}\right)f\left(x+\frac{y}{2},p\right)\right\}\times$$

$$\times\exp\left(-2\frac{i}{\hbar}py\right)dx\,dp$$

$$= \int x_i f(x,p)\exp\left(-2\frac{i}{\hbar}py\right)dx\,dp. \qquad (7\cdot9)$$

In similar fashion, we multiply (7.7) successively with $x_i\,x_j,\ x_i\,x_j\,x_k,\cdots$, and then integrate the results over the whole x-space to get a series of relations involving p-differentiation of successively higher orders. This series of relations, (7.8), $(7.9),\cdots$, may be regarded as equivalent with the original (7.7).

The second step is the Taylor expansions of (7.8), (7.9),⋯ with respect to y. Then we get from (7.8) the relations :

$(7.2°)$,

$$\hbar^3 \int p_i f^2 \, dx \, dp = \int p_i f \, dx \, dp, \qquad (7 \cdot 10)$$

$$\hbar^3 \int \left(\frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_k} + \frac{4}{\hbar^2} p_i p_k f^2 \right) dx \, dp = \frac{4}{\hbar^2} \int p_i p_k f \, dx \, dp, \qquad (7 \cdot 11)$$

$\cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots$

Similarly (7.9) yields

$$\hbar^3 \int x_i f^2 \, dx \, dp = \int x_i f \, dx \, dp, \qquad (7 \cdot 12)$$

$$\hbar^3 \int \left(-\frac{\partial f}{\partial x_k} \frac{\partial f}{\partial p_i} + \frac{4}{\hbar^2} x_i p_k f^2 \right) dx \, dp = \frac{4}{\hbar^2} \int x_i p_k f \, dx \, dp, \qquad (7 \cdot 13)$$

$\cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots$

We have thus found that the pure state condition can also be expressed with a set of infinitely many relations,

$$(7.2°), \quad (7.10), \quad (7.11), \cdots; \quad (7.12), \quad (7.13), \cdots; \cdots,$$

which are symmetrical in x and p and of different expressions from the previously obtained ones,

$$(4.6), \quad (4.19); \quad \text{or} \quad (4.19), \quad (4.18), \quad (4.21), \cdots.$$

But both forms should be equivalent with each other, which fact may be understood if we return to the expressions in ρ, though it may be difficult to show it directly.*

(c) We could take either (4.2) or (4.4) as the pure state condition in terms of density matrix. We have, however, still another relation valid for a pure state,

$$\rho(xx) \rho(x'x') = \rho(xx') \rho(x'x) = |\rho(xx')|^2, \qquad (7 \cdot 14)$$

though it is a necessary but not sufficient relation for the pure state condition. It can be shown that the relation (7.14) is derivable from (4.4) by the Taylor expansion of the former. It is also to be noted that the integration of (7·14) throughout over the whole x'-space yields

$$\rho(xx) = \int \rho(xx') \rho(x'x) dx, \qquad (7 \cdot 15)$$

which is also a special case of (4.2).

Now we shall translate the relation (7.14) into the language of the ps. df. to obtain

$$P(x - y/2) \, P(x + y/2) = \left| \int f(x, p) e^{-ipy/\hbar} \, dp \right|^2. \qquad (7 \cdot 16)$$

_____

*) For this point, see also (c).

If we next expand (7.16) into the power series of γ', (7.16) turns into a set of relations,

$$\sum_{\substack{r_1 r_2 r_3 \ s_1 s_2 s_3 \\ (r_i + s_i = n_i)}} \frac{(-1)^{\frac{\Sigma r_i}{i}}}{r_1!\, r_2!\, r_3!\, s_1!\, s_2!\, s_3!} [P_{r_1 r_2 r_3}(\boldsymbol{x}) P_{s_1 s_2 s_3}(\boldsymbol{x})$$

$$- (-\varepsilon)^{n/2} \partial_1{}^{r_1} \partial_2{}^{r_2} \partial_3{}^{r_3} P(\boldsymbol{x}) \cdot \partial_1{}^{s_1} \partial_2{}^{s_2} \partial_3{}^{s_3} P(\boldsymbol{x})] = 0,$$

$$(n \equiv \sum_i n_i = 2,\ 4, \cdots\cdots), \qquad\qquad (7\cdot17)$$

with

$$P_{r_1 r_2 r_3} = \int p_1{}^{r_1} p_2{}^{r_2} p_3{}^{r_3} f(\boldsymbol{x},\, \boldsymbol{p})\ d\boldsymbol{p},$$

where the summation is to be taken for zero or integer values of each r_i, s_i, with $r_i + s_i = n_i$ fixed. The lowest relation ($n=2$) of (7.17) is nothing but (4.13) (i.e.,(4.18)), and the next one involves the fourth order moment.* Thus we see that the relation (7.17) or (7.16) consists of just a *half* of the former relations (4.19), (4.18), (4.21),⋯ of the pure state condition corresponding to (4.4).

The relation (7.15) is expressed in terms of f as (7.5) with equality sign, which can be split into a series of relations, (7.2°), (7.12),⋯ by the procedure of taking moments with respect to x. These relations, on one hand, naturally coincide with a part of relations obtained in (b), and, on the other, must be derivable from (7.17), since (7.15) was a result of (7.14).

Our method is essentially the transcription of von Neumann's density matrix method[8a] through a particular fourier transformation of a specified representation of the latter which formulates quantum mechanics generally for mixture and characterizes a pure state by a subsidiary condition, thus leading to formulating quantum mechanics in a closed form in such a way that it is associated with the phase-space ensemble picture.

Really the method was described for the case of single non-relativistic particle, but it would also be interesting to apply this method to the case of a many particle system.

---

*) For one-dimensional case it is

$$P P_4 - 4 P_1 P_3 + 3 P_2{}^2 = (\hbar/4)^4 \{ P \partial^4 P - 4 \partial P \cdot \partial^3 P + 3 (\partial^2 P)^2 \},$$

with
$$P_r = \int p^r f(xp)\ dp.$$

Appendix

A. Examples of the momentum transition probability $J(x, p)$, simple but not reducible to $\nabla V \, \nabla_p \delta(p)$

i)   For a linear rectangular potential barrier : $V(x) = V_0$ for $|x| \lesssim a/2$, and 0 for $|x| \gtrsim a/2$; we have $V(p) = 1/\pi \cdot V_0 \sin(pa/2\hbar)/p$, and so

$$J(x, p) = -2V_0/\pi\hbar p \cdot \sin(pa/\hbar) \sin(2px/\hbar).$$

ii)   For a periodic potential, $V(x) = \sum_n V_n e^{ik_n x}$, we have $V(p) = \sum_n V_n \delta(p + \hbar k_n)$, and so

$$J(x, p) = -\frac{2}{\hbar} \, \mathrm{Im}\Big[\sum_n V_n \, e^{2ipx/\hbar}\Big] \cdot \delta\Big(p + \frac{1}{2}\hbar k_n\Big).$$

Therefore particle momentum can jump only by an amount of some 'resonance value' $\hbar k_n/2$.

iii)   For Coulomb potential $V = -e^2/r$, we have $V(p) = -(e^2/2\pi^2\hbar) \, 1/p^2$, so we obtain

$$J(x, p) = \frac{2e^2}{\pi^2\hbar^2} \, \frac{\sin(2px/\hbar)}{p^2}.$$

In this case momentum cannot jump in the direction perpendicular to the force.

B. Alternative procedure for obtaining the pure state conditions in the form of § 4(b)

When our pure state condition for the density matrix, (4.4), is satisfied, we can derive from it the relations,

$$D(n_1 n_2 n_3)\rho \cdot \partial_k' \rho - \rho \cdot D(n_1 n_2 n_3)\partial_k' \rho = 0, \qquad \qquad (B \cdot 1)$$

by differentiation, or further, even more general form

$$D(n_1 n_2 n_3)\rho \cdot D'(n_1' n_2' n_3')\rho - \rho \cdot D(n_1 n_2 n_3) D'(n_1' n_2' n_3')\rho = 0, \qquad (B \cdot 2)$$

where $D(n_1 n_2 n_3) \equiv \partial_1^{n_1} \partial_2^{n_2} \partial_3^{n_3}$, $D'(n_1' n_2' n_3') \equiv \partial_1'^{n_1'} \partial_2'^{n_2'} \partial_3'^{n_3'}$ with $\partial_i = \partial/\partial x_i$, $\partial_k' = \partial/\partial x_k'$, n_i and n_i' being non-negative integers[*].

Now, if (4.4) holds for any values of x and x', we must have

$$[D(n_1 n_2 n_3)(\partial_i \rho \partial_k' \rho - \rho \partial_i \partial_k' \rho)]_{x=x'} = 0, \qquad \qquad (B \cdot 3)$$

$$\text{for } \sum_i n_i = 0, 1, 2, \cdots \infty,$$

by the Taylor expansion of (4.4) on the diagonal.   (B.3) can further by replaced by

$$[D(n_1 n_2 n_3)\rho \cdot \partial_k' \rho - \rho \cdot D(n_1 n_2 n_3)\partial_k' \rho]_{x=x'} = 0, \qquad \qquad (B \cdot 4)$$

$$\text{for } \sum_i n_i = 1, 2, 3, \cdots \infty.$$

---

[*]   It is to be noted that the compatibility of the pure state condition (4.4) with the equation of motion (3.2) for ρ can here be shown explicitly, as we can derive

$$\partial/\partial t \, (\partial_i \rho \partial_k' \rho - \partial_i \partial_k' \rho) = 0$$

from (3.2) and (B.2).

We can thus replace the pure state condition (4.4) by *a series of relations on the diagonal*, (B.4). The lowest relation of (B.4) is

$$[\partial_i \rho \, \partial_k' \rho - \rho \, \partial_i \, \partial_k' \rho]_{x=x'} = 0, \tag{B·5}$$

and its complex conjugate is

$$[\partial_k \rho \, \partial_i' \rho - \rho \, \partial_k \, \partial_i' \rho]_{x=x'} = 0, \tag{B·6}$$

which is, however, identical with the relation (B.5) with i and k interchanged. Therefore the nine *complex* relations, (B.5), are not independent of each other but consist of six symmetrized *real* equations,

$$[(\partial_i \rho \, \partial_k' \rho + \partial_k \rho \, \partial_i' \rho) - \rho(\partial_i \, \partial_k' \rho + \partial_k \, \partial_i' \rho)]_{x=x'} = 0, \tag{B·7}$$

and three anti-symmetrized *real* ones,

$$i[(\partial_i \rho \, \partial_k' \rho - \partial_k \rho \, \partial_i' \rho) - \rho(\partial_i \, \partial_k' \rho - \partial_k \, \partial_i' \rho)]_{x=x'} = 0. \tag{B·8}$$

The next relations of (B.4),

$$[\partial_i \, \partial_j \rho \, \partial_k' \rho - \rho \, \partial_i \, \partial_j \, \partial_k' \rho]_{x=x'} = 0 \tag{B.9}$$

are not independent among themselves, too, on account of (B.5).

Now our next task is to represent the conditions (B.4) in terms of the distribution function. First we re-express (B.4) in terms of $\bar{\rho}(x, y)$ to obtain

$$\left[\left\{ \prod_{l=1,2,3} \left(\frac{1}{2} \frac{\partial}{\partial x_l} - \frac{\partial}{\partial y_l} \right)^{n_l} \right\} \bar{\rho} \cdot \left(\frac{1}{2} \frac{\partial}{\partial x_k} + \frac{\partial}{\partial y_k} \right) \bar{\rho} \right]_{y=0}$$

$$= \left[\bar{\rho} \cdot \left\{ \prod_{l=1,2,3} \left(\frac{1}{2} \frac{\partial}{\partial x_l} - \frac{\partial}{\partial y_l} \right)^{n_l} \right\} \left(\frac{1}{2} \frac{\partial}{\partial x_k} + \frac{\partial}{\partial y_k} \right) \bar{\rho} \right]_{y=0}. \tag{B·10}$$

Here, using the relation

$$\left[\left\{ \prod_l \left(\frac{\partial}{\partial y_l} \right)^{r_l} \right\} \bar{\rho} \right]_{y=0} = \left(-\frac{i}{\hbar} \right)^{r_1+r_2+r_3} P_{r_1 r_2 r_3}(x),$$

with

$$P_{r_1 r_2 r_3} \equiv \int p_1^{r_1} p_2^{r_2} p_3^{r_3} f(x, p) \, dp,$$

(B.10) is transformed into the relation :

$$\sum_{r_1, r_2, r_3=0}^{n_1, n_2, n_3} \left(\frac{2i}{\hbar} \right)^{r_1+r_2+r_3} \binom{n_1}{r_1}\binom{n_2}{r_2}\binom{n_3}{r_3} \left[\left(\prod_l \partial_l^{n_l-r_l} \right) P_{r_1 r_2 r_3} \left(\frac{1}{2} \partial_k P - \frac{i}{\hbar} P_k \right) \right.$$

$$\left. - P \left(\prod_l \partial_l^{n_l-r_l} \right) \left(\frac{1}{2} \partial_k P_{r_1 r_2 r_3} - \frac{i}{\hbar} P_{(r_1 r_2 r_3)} + 1_k \right) \right] = 0, \tag{B·11}*$$

which may further be separated into real and imaginary part equations. Thus we have obtained the pure state conditions in the form of relations between distribution moments

---

*) $P_{(r_1 r_2 r_3)+1_l} \equiv P_{r_1+\delta_{1,k}, \ r_2+\delta_{2,k}, \ r_3+\delta_{3,k}}$

of successively higher orders in respect to p. Especially we find that the lowest relations (B. 7) and (B. 8) transform, according to these procedures, just into (4.18) and (4.19), respectively, while the transforms of the next relation (B.9) become (4.21).

References

1) The content of this article was reported in Prog. Theor, Phys. **10** (1953), 119, 121, 122, in its essentials, and also in Busseiron-Kenkyu No. 66 (Sept. 1953), p. 48, in Japanese nearly in full .

2) P. A. M. Dirac, *The Principles of Quantum Mechanics*, third ed. p. 10.

3) R. P. Feynman, Rev. Mod. Phys. **20** (1948), 367.

4) T. Takabayasi, Prog. Theor. Phys. 8(1952), 143; also see D. Bohm, Phys. Rev. **84** (1952), 166, 180; L. de Broglie, *La Physique Quantique Restra-t-elle Indéterministe ?* (Gauthier-Villars, Paris, 1953).

5) T. Takabayasi, Prog. Theor. Phys. **9** (1953), 187.

6) E. P. Wigner, Phys. Rev. **40** (1932), 749.

7) J. E. Moyal, Proc. Cambridge Phil. Soc. **45** (1949), 99.

8) H. S. Green, J. Chem. Phys. **19** (1951), 955; J. H. Irving and R. W. Zwanzig, J. Chem. Phys. **19** (1951), 1173; H. Mori and S. Ono, Prog. Theor. Phys. 8 (1952), 327; H. Mori, ibid. 9(1953), 473.

8a) J. v. Neumann, *Mathematische Grundlagen der Quanteumechanik* (1932).

9) H. Weyl, *Gruppentheorie und Quantenmechanik* (1931), p. 244.

10) Cf. P.A. M. Dirac, Rev. Mod. Phys. **17** (1945), 195.

11) Cf. F. Bopp, ZS. f. Naturforsch. 2a (1947), 202; also cf. reference 4) §1.

12) See T. Takabayasi, J. of History of Science, Japan, No. 19 (1951). It would be interesting to compare our viewpoint with that of K. Husimi, Prog. Theor. Phys. 9(1953), 381.

13) H. A. Kramers, Physica 7(1940), 284.

14) W. Weizel, ZS. f. Phys. **134** (1953), 264.

15) Compare with D. Bohm, *Quantum Theory* (Prentice-Hall, New York, 1951) p. 240.

16) Cf. E. Schrödinger, Naturwiss. **14** (1926), 664; L. Schiff, *Quantum Mechanics*, p. 67.

Note added in proof

1) To page 345 : For instance, the q-no. function due to Weyl's procedure corresponding to a c-no. function $A(x, p) = x_i p_i$ is found to be $(x_i p_i + p_i x_i)/2$, so

$$<(x_i p_i + p_i x_i)/2>_\text{qu} = \int x_i p_i f(x, p) \ dx \ dp.$$

Further, if we take this hermitian quantity not belonging to (2.12b) for the A in (2·3), we obtain another one of positivity conditions,

$$\int x_i^2 p_i^2 f(x, p) \ dx \ dp \geq -\hbar^2/4.$$

2) To page 346 : A quantum-mechanical state compatible with given $P(x)$ and $Q(p)$ cannot be determined uniquely even when the state is restricted within pure cases (see K. Husimi, Kagaku 5 (1935), 370).

3) To page 352 : The transformation function T for the case of forced oscillator was given by T. Nishiyama, Prog. Theor. Phys. 8 (1952), 655.

4) To § 5 (c): Recently Weizel applied his method to many particle system (ZS. f. Phys. **136** (1954), 582).

5) To § 7 (a): The positivity condition is to be regarded as involving the *uncertainty* for the case of mixture, since it means that a permissible ps. df. $f(x, p)$ should be expanded as the superposition with positive coefficients of mutually orthogonal ps. dfs. f_n satisfying pure state conditions, as

$$f(x, p) = \sum_n w_n f_n(x, p),$$

with $w_n \geq 0$, $\int f_n(x, p) f_m(x, p) \ dx \ dp = \delta_{nm}/h^3$. It was, however, not easy to express the sufficient positivity condition in a more convenient form.

A Method in Quantum Statistical Mechanics, I
—— *General Theory* ——

Hiroshi ICHIMURA

Physics Department, Tokyo Institute of Technology
Oh-Okayama, Meguro-ku, Tokyo

A formal expansion method for the grand partition function of the system of interacting particles is developed, in which the number representation and the ordered exponential method are used. The resulting formula will be useful in calculating quantum effects-statistical and dynamical- in the thermodynamical property of various systems. Applications to the special problems will be given in later papers.

Introduction

In these two decades, many authors[1-6] have put forward various methods for the calculation of the partition function of the system of the interacting particles on the basis of the quantum theory. Although their method of attacks have their own characteristic features, most of them use the density matrix in the ordinary q-representation, and frequently the molecular distribution functions play important roles. It seem to us, however, the use of the q-representation is not always convenient.

Here we want to make an attempt to use the number representation of the second quantization theory for the calculation of the trace of the density matrix. In this representation, we think, the effects of the quantum theory may most easily be taken into account. In Part I we derive an expansion formula for the grand partition function in the form of the product of the grand partition function of the ideal gas and the correction factor due to the interaction between particles. Since we have not succeeded in deriving the general term of the expansion, the discussion of the convergency of our procedures must be done on each special occasion. In Part II and III which will soon appear we treat the highly degenerated Fermi assembly and the virial expansion for the quantum gases. In the latter case we obtain the formula of the second virial coefficient which is effective at low temperatures.

Part I. General Theory

§ 1. Grand partition function in the number representation

Following the usual procedure[7] we can express the Hamiltonian of the system of identical particles as

$$H = H^{(0)} + U,$$

$$H^{(0)} = \sum_r \mathcal{E}_r a_r^+ a_r , \quad \mathcal{E}_r = \hbar^2 k_r^2/2m , \quad k_r = |\boldsymbol{k}_r|,$$

$$U = \frac{J}{4} \sum_p \sum_q \sum_r \sum_s a_p^+ a_q^+ a_s a_r (G_{pq,rs}),$$

$$J(G_{pq,rs}) = JG_{pq,rs} \mp JG_{pq,sr}$$

$$(- \text{Fermi}, + \text{Bose}), \tag{1.1}$$

$$JG_{pq,rs} = \frac{\Theta_{pq,rs}}{V^2} \iint G(r_{12}) \exp[i(\boldsymbol{k}_r - \boldsymbol{k}_p)\boldsymbol{r}_1] \exp[i(\boldsymbol{k}_s - \boldsymbol{k}_q)\boldsymbol{r}_2] d\boldsymbol{r}_1 d\boldsymbol{r}_2,$$

$$\Theta_{pq,rs} = \theta_p(\xi_1)\theta_q(\xi_2)\theta_r(\xi_1)\theta_s(\xi_2),$$

$$r_{12} = |\boldsymbol{r}_2 - \boldsymbol{r}_1|.$$

Here we have taken as the basic orthogonal system of functions the plane waves normalized in the volume V, i.e.

$$\frac{1}{\sqrt{V}} e^{ikr};$$

$JG(r_{12})$ means the interaction potential between two particles which lie at $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$; $\theta_p(\xi)$ is the normalized spin eigenfunction for the p-th state; the summations extend over the momentum and spin states. Because of our special formalism, we must be careful for the nature of the singularity of $G(r)$ at $r=0$ and $r=\infty$. In the following we use only the Morse type functions. We cannot adopt the Coulomb potential and the rigid sphere potential owing to their singular natures.

a^+ and a are operators which usually appear in the second quantization theory and have the following properties.

Fermi-Dirac Statistics

$$a_r^+ a_r = n_r , \qquad a_r a_r^+ = 1 - n_r,$$

$$a_i a_k^+ + a_k^+ a_i = \delta_{ik}. \tag{1.2}$$

Bose-Einstein Statistics (We use b^+ and b in this case)

$$b_r^+ b_r = n_r, \qquad b_r b_r^+ = 1 + n_r,$$

$$b_i b_k^+ - b_k^+ b_i = \delta_{ik}, \tag{1.3}$$

n_r means the operator which expresses the number of particles in r-th state. Concerning the results of operations of these operators on the wave functions in the number representation, we follow the usages in the reference (8).

As our basic system of the representation we take the eigen states of $H^{(0)}$. This is always permitted on account of the invariant character of the trace under the unitary transformation.

Using the above mentioned representation, we can define the grand partition function

$Z(\lambda, \beta)$ for the system of interacting particles as

$$\dot{Z}(\lambda, \beta) = S\rho(\lambda^{\Sigma n_r} \exp\left[-\beta(\boldsymbol{H}^{(0)} + U)\right]). \tag{1.4}$$

Here $S\rho$ means to take the trace of the operator in the bracket. In our basic system for which all n's are diagonal, this $S\rho$ is equivalent to the multiple summation

$$\sum_{n_1} \sum_{n_2} \cdots \sum_{n_r} \cdots,$$

where all n_r's take the values 0, 1 in Fermi-Dirac case and 0, 1, 2, 3, $\cdots\cdots$ in Bose-Einstein case. From the usual relations such as

$$\bar{N} = \lambda \partial \log Z / \partial \lambda,$$

$$pV = \kappa T \log Z, \tag{1.5}$$

$$\lambda = \exp\left[\beta\mu\right], \quad \beta = 1/\kappa T,$$

we can determine the thermodynamical properties of our system. Here $\bar{N}$ is the average number of particles in the volume V, and μ is the thermodynamical potential per particle and κ is the Boltzmann constant.

In (1.4), we can divide U into two parts D and N; D is the diagonal part and N is the off-diagonal part.

$$JD = J \sum_{j \leqslant k} \sum n_j n_k G_{jk}, \tag{1.6}$$

$$JG_{jk} = JG_{jk,jk} \mp JG_{jk,kj}, \qquad (-\text{Fermi}, +\text{Bose}) \tag{1.7}$$

$$JN = J \sum_{p \leqslant q} \sum_{(\neq)} \sum_{r \leqslant s} \sum a_p^+ a_q^+ a_s a_r (G_{pq,rs}) \tag{1.8}$$

$$= J S_{pq,rs} a_p^+ a_q^+ a_s a_r (G_{pq,rs}). \tag{1.8'}$$

In (1.8), the symbol $\neq$ means that we should exclude the terms in which $a_p^+ a_q^+ a_s a_r$ are the diagonal forms; the four fold summation reduces to the three-fold one because of the condition of non-zero $(G_{pq,rs})$, as easily be seen from the definition in (1.1)

$$\boldsymbol{k}_r + \boldsymbol{k}_s = \boldsymbol{k}_p + \boldsymbol{k}_q. \tag{1.9}$$

To indicate these restrictions, we have used the special symbol $S_{pq,rs}$ in (1.8').

Now we use the method invented by Feynman for the expansion of the operator

$$\exp\left[-\beta(\boldsymbol{H}^{(0)} + U)\right] \tag{1.10}$$

in powers of U, where $\boldsymbol{H}^{(0)}$ and U are incommutative. Then, in our representation scheme, we can express the $\nu\nu$-element of the above operator as follows, where $H_\nu^{(0)}$ and D_ν mean the ν-th diagonal element:

$$(\exp\left[-\beta(\boldsymbol{H}^{(0)} + U)\right])_{\nu\nu} = (O_0)_{\nu\nu} + (O_1)_{\nu\nu} + (O_2)_{\nu\nu} + \cdots,$$

$$(O_0)_{\nu\nu} = \exp\left[-\beta H_\nu^{(0)}\right], \qquad H^{(0)} = \sum_r n_r \varepsilon_r,$$

$$(O_1)_{\nu\nu} = -\beta J \exp\left[-\beta H_\nu^{(0)}\right] D_\nu,$$

$$(O_2)_{\nu\nu} = \beta^2 J^2 \left\{ \frac{1}{2!} \exp\left[-\beta H_\nu^{(0)}\right] D_\nu^2 + \sum_{\sigma(\neq\nu)} \left(\frac{e^{-\beta H_\nu^{(0)}}}{\beta(H_\sigma^{(0)} - H_\nu^{(0)})} \right. \right. \tag{1.11}$$

$$\left. \left. + \frac{e^{-\beta H_\sigma^{(0)}} - e^{-\beta H_\nu^{(0)}}}{\beta^2(H_\sigma^{(0)} - H_\nu^{(0)})^2} \right) N_{\nu\sigma} N_{\sigma\nu} \right\}.$$

. . . .

Further formulas are given in the appendix.

Using this formula, we obtain an expansion of Z in powers of J;

$$Z = \sum_i Z_i,$$

$$Z_i = Sp(\lambda^{\sum n_r} O_i). \tag{1.12}$$

In the following section, we calculate each terms in (1.12) and show that Z can be approximated in the form

$$Z = Z_0 \exp(S),$$

$$S = J C_1 + J^2 C_2 + \cdots = \sum_i J^i C_i. \tag{1.13}$$

Here

$$Z_0 = \prod_r (1 \pm \lambda e^{-\beta\varepsilon})^{\pm 1}$$

$$(+\,\text{Fermi}, \ -\,\text{Bose}) \tag{1.14}$$

is the grand partition function for the system of non-interacting particles.

§ 2. Calculation of Z_i

(1) Z_0: It is rather obvious that this term expresses the grand partition function of the ideal gas system.

$$Z_0 = Sp(\lambda^{\sum n_r} \exp\left[-\beta \sum_r n_r \varepsilon_r\right]) = Sp(\prod_r (\lambda e^{-\beta\varepsilon_r})^{n_r},\ )$$

$$= \prod_r (1 \pm \lambda e^{-\beta\varepsilon_r})^{\pm 1}, \tag{2.1}$$

$$(+\,\text{Fermi}, \ -\,\text{Bose}).$$

(2) Z_1: From the equations (1.11) and (1.7) we get

$$Z_1 = -\beta J Sp(\lambda^{\sum n_r} \exp\left[-\beta \sum_r \varepsilon_r n_r\right] (\sum_j \sum_{\leq k} n_j n_k G_{jk}))$$

$$= -\beta J Z_0 \langle D \rangle$$

$$= Z_0 J D_1, \tag{2.2}$$

where

$$J D_1 = -\beta J \sum_j \sum_{\leq k} f_j f_k G_{jk}, \tag{2.3}$$

$$f_j = \frac{\lambda e^{-\beta\varepsilon_j}}{1 \pm \lambda e^{-\beta\varepsilon_j}}, \tag{2.4}$$

$$(+\,\text{Fermi}, \ -\,\text{Bose}).$$

This is trivial for the Fermi case and we use the relation

$$\lambda \frac{\partial}{\partial \lambda} \frac{1}{(1-\lambda e^{-\beta \varepsilon})} = \frac{\lambda e^{-\beta \varepsilon}}{(1-\lambda e^{-\beta \varepsilon})^2} = \lambda e^{-\beta \varepsilon} + 2\lambda^2 e^{-2\beta \varepsilon} + \cdots$$

for the Bose case.

(3) Z_2: From $(1 \cdot 11)$, $(1 \cdot 7)$ and $(1 \cdot 8)$, it becomes

$$Z_2 = Z_2^{(1)} + Z_2^{(2)},$$

$$Z_2^{(1)} = \beta^2 J^2 Sp \left(\prod_r (\lambda e^{-\beta \varepsilon_r})^{n_r} \frac{1}{2!} D^2 \right),$$

$$Z_2^{(2)} = \beta^2 J^2 Sp \left(\lambda^{\Sigma n_r} \sum_{\sigma(\neq \nu)} \frac{e^{-\beta H_\nu^{(0)}}}{\beta(H_\sigma^{(0)} - H_\nu^{(0)})} \right). \tag{2.5}$$

The contribution from the last term in O_2 vanishes (see Appendix).

In the same way as in (2), we get

$$Z_2^{(1)} = \frac{\beta^2 J^2}{2!} Sp \left(\prod_r (\lambda e^{-\beta \varepsilon_r})^{n_r} \left(\sum_j \sum_{\geq k} n_j n_k G_{jk} \right) \left(\sum_l \sum_m n_l n_m G_{lm} \right) \right)$$

$$= Z_0 \frac{\beta^2 J^2}{2!} \langle D^2 \rangle$$

$$= Z_0 \left(\frac{J^2}{2!} D_1^2 + J^2 D_2 + J^2 R_2 \right), \tag{2.6}$$

$$J^2 D_2 = \frac{\beta^2 J^2}{3!} \sum_{j \neq k \neq l} f_j f_k f_l$$

$$\times \{ (1 \mp f_j) G_{jk} G_{jl} + (1 \mp f_k) G_{jk} G_{kl} + (1 \mp f_l) G_{jl} G_{kl} \}, \tag{2.7}$$

$$(-\text{Fermi}, +\text{Bose}).$$

We do not give the explicit form of R_2, since the contribution from this term is neglected in the final results as we shall show later.

For the calculation of the second term in $(2 \cdot 5)$, we need the non-vanishing diagonal element $J^2 N_{\nu \sigma} N_{\sigma \nu}$. Here as we know

$$\nu = (n_1, n_2, \cdots),$$

$$\sigma = (n_1', n_2', \cdots).$$

Then from the definition in $(1 \cdot 8)$, the non-vanishing term will be given by

$$J^2 n_{pq, rs} = J^2 (1-n_p)(1-n_q) n_r n_s |(G_{pq, rs})|^2, \qquad (\text{Fermi}), \tag{2.8}$$

$$J^2 n_{pq, rs} = J^2 (n_p+1)(n_q+1+\delta_{pq}) n_r (n_s-\delta_{rs}) |(G_{pq, rs})|^2, \quad (\text{Bose}) \tag{2.8}$$

and the summation $\sum_m$ with fixed ν is equivalent to the summation

$$S_{pq, rs}$$

and further

$$H_\sigma^{(0)} - H_\nu^{(0)} = \mathcal{E}_{pq} - \mathcal{E}_{rs},$$

$$\mathcal{E}_{xy} = \mathcal{E}_x + \mathcal{E}_y \tag{2.9}$$

under this summation.

Hence we have

$$Z_2^{(2)} = \beta^2 J^2 Sp\left(\prod_r (\lambda e^{-\beta \varepsilon_r})^{n_r} S_{pq, rs} \frac{n_{pq, rs}}{\beta(\mathcal{E}_{pq} - \mathcal{E}_{rs})} \right)$$

$$= Z_0(J^2 \langle N_2 \rangle), \tag{2.10}$$

where

$$N_2 = \beta^2 S_{pq, rs} \frac{n_{pq, rs}}{\beta(\mathcal{E}_{pq} - \mathcal{E}_{rs})}$$

and

$$\langle N_2 \rangle = \beta^2 S_{pq, rs} \frac{|(G_{pq, rs})|^2}{\beta(\mathcal{E}_{pq} - \mathcal{E}_{rs})} (1 \mp f_p)(1 \mp f_q) f_r f_s, \tag{2.11}$$

$$(-\text{Fermi}, +\text{Bose}).$$

(4) Z_3 : In similar fashions, we can proceed as follows:

$$Z_3 = Z_3^{(1)} + Z_3^{(2)},$$

$$Z_3^{(1)} = -\beta^3 J^3 S_p\left(\prod_r (\lambda e^{-\beta \varepsilon_r})^{n_r} \frac{1}{3!} D^3 \right),$$

$$Z_3^{(2)} = -\beta^3 J^3 S_p(\prod_r (\lambda e^{-\beta \varepsilon_r})^{n_r} F_3), \tag{2.12}$$

where F_3 represents all the terms in (A·10) except the first. $Z_3^{(1)}$ can be transformed as

$$Z_3^{(1)} = Z_0 \frac{J^3}{3!} \langle D^3 \rangle$$

$$= Z_0 \left(\frac{J^3}{3!} D_1^3 + J^3 D_1 D_2 + J^3 D_3 + J^3 R_3 \right), \tag{2.13}$$

where

$$J^3 D_3 = \frac{1}{4!} \beta^3 J^3 \sum_{j \neq k \neq l \neq q} f_j f_k f_l f_q$$

$$\times [\{jkl, q\} + \{klq, j\} + \{lqj, k\} + \{qjk, l\}]$$

$$\{jkl, q\} = \{(1 \mp f_j) G_{jk} G_{jl} + (1 \mp f_k) G_{jk} G_{kl} + (1 \mp f_l) G_{jl} G_{kl}\}$$

$$\times \{(1 \mp f_j) G_{jq} + (1 \mp f_k) G_{kq} + (1 \mp f_l) G_{lq}\}, \tag{2.14}$$

$$(-\text{Fermi}, +\text{Bose}).$$

---

* We omit the terms which give the contribution of order in the final results.

Again R_3 is the redundant terms. Now we can write for

$$Z_3 = Z_0 \left(\frac{J^3}{3!} \langle D^3 \rangle + J^3 \langle F_3 \rangle \right). \tag{2·15}$$

For Z_4, Z_5, $\cdots\cdots$, we can also write

$$Z_n = Z_0 \left(\frac{J^n}{n!} \langle D^n \rangle + J^n \langle F_n \rangle \right). \tag{2·16}$$

§ 3.　Determination of C_i

The formal expansion of $\exp\left[\sum J^i C_i\right]$ has the form

$$\exp\left[\sum J^i C_i\right] = 1 + J C_1 + J^2 \left(C_2 + \frac{1}{2!} C_1^2 \right)$$

$$+ J^3 \left(C_3 + C_1 C_2 + \frac{1}{3!} C_1^3 \right) +$$

$$+ J^4 \left(C_4 + C_1 C_3 + \frac{1}{2!} (C_2)^2 + \frac{1}{2!} (C_1)^2 C_2 + \frac{1}{4!} C_1^4 \right)$$

$$+ \cdots. \tag{3·1}$$

Considering this form and the thermodynamical requirement that $\log Z$ should have the functional dependence

$$\bar{N} f(\bar{N}/V) \tag{3·2}$$

on $\bar{N}$, we rearrange the results in § 2 as follows.

We define C_i's in terms of D_i and others successively,

$$C_1 = D_1,$$

$$C_2 = \langle\!\langle \frac{1}{2!} (\langle D^2 \rangle - \langle D \rangle^2) + \langle N_2 \rangle \rangle\!\rangle = D_2 + \langle N_2 \rangle,$$

$$C_3 = \langle\!\langle \frac{1}{3!} \langle D^3 \rangle + \langle F_3 \rangle - \left(C_1 C_2 + \frac{1}{3!} C_1^3 \right) \rangle\!\rangle$$

$$= D_3 + \langle\!\langle \langle F_3 \rangle - \langle D \rangle \langle N_2 \rangle \rangle\!\rangle,$$

$$C_4 = \langle\!\langle \frac{1}{4!} \langle D^4 \rangle + \langle F_4 \rangle - \left(C_1 \cdot C_3 + \frac{1}{2!} C_2^2 + \frac{1}{2!} C_1^2 C_2 + \frac{1}{4!} C_1^4 \right) \rangle\!\rangle \tag{3·3}$$

$$= D_4 + \langle\!\langle \langle F_4 \rangle - \left(\langle D \rangle \langle\!\langle \langle F_3 \rangle - \langle D \rangle \langle N_2 \rangle \rangle\!\rangle + \frac{1}{2!} \langle D \rangle^2 \langle N_2 \rangle + D_2 \langle N_2 \rangle + \frac{1}{2!} \langle N_2 \rangle^2 \right) \rangle\!\rangle,$$

$$D_4 = \langle\!\langle \frac{1}{4!} \langle D^4 \rangle - \left(D_1 \cdot D_3 + \frac{1}{2!} D_2^2 + \frac{1}{2!} D_1^2 D_2 + \frac{1}{4!} D_1^4 \right) \rangle\!\rangle,$$

$$\cdots\cdots\cdots\cdots.$$

The symbol $\langle\!\langle\ \ \rangle\!\rangle$ means to take only those terms of the order $\overline{N}$, that is

$$\langle F_3 \rangle - \langle D \rangle \langle N_2 \rangle = \langle\!\langle \langle F_3 \rangle - \langle D \rangle \langle N_2 \rangle \rangle\!\rangle + R_3'.$$

We can see that $D_1,\ D_2,\ \cdots$ are of order $\overline{N}$.

Then we have

$$Z = Z_0 \left(\exp\left[\sum J^i C_i\right] + J^3 (R_3 + R_3') + J^4 (R_4 + R_4') \cdots \right). \tag{3.4}$$

In general, R_3 may have the order $\overline{N}^2$ and R_4 the order $\overline{N}^3, \cdots$.

Rearranging formally, we have the form

$$Z = Z_0 \exp\left[\sum_i J^i C_i\right](1 + \overline{N}a + \overline{N}^2 b + \cdots),$$

where we have indicated explicitly the orders in $\overline{N}$ of each term.

From the thermodynamical requirement, the $\overline{N}$ dependence of the last factor will be at most

$$\exp\left(\overline{N}\sum J^k \delta_k\left(\frac{\overline{N}}{V}\right)\right),$$

but in this case, this factor has to be included in $\exp\left[\sum J^i C_i\right]$ from the beginning. Therefore we can expect that the last factor which comes from the redundant terms has the weaker $\overline{N}$ dependence. So the contributions from this factor in $\log Z$ are negligible compared to $\overline{N}$.

Hence we can approximate our grand partition function in the form

$$Z = Z_0 \exp\left[\sum_i J^i C_i\right]$$

with C_i's determined by the relations in (3.3).

§ 4. Conclusion

Our series in (1.13) is the expansion in powers of the interaction parameter J and as we have not assumed anything on the state of agregation of our system, it may be used in both gas and liquid phases. But the convergency of the series should be examined carefully on each special case. If the series converges our formula will offer a rather simple method to treat the quantum effects (statistical and dynamical) in the thermodynamical property of the system under consideration. In the next papers, we will treat its application to the highly degenerated Fermi gas in which λ is very large and to the virial expansion for the quantum gases in which λ is small. And further we want to treat the problem of the interaction between the system of conduction electrons and the lattice vibrations[11] by the present method.

Appndix. Derivation of (1.11)

In his paper on the quantum electrodynamics, Feynman introduced a formal method to expand the operator $\exp[A+B]$ in powers of the operator B, where A and B do not commute, and the operator $\exp[x]$ is defined as $\sum_0^\infty \frac{1}{n!}x^n$.

In his method, it is written as

$$\exp[A+B] = \exp[\int_0^1 (A_s + B_s)\,ds].\tag{A$\cdot$1}$$

The suffixs s on A_s or B_s indicates the order of the operation in the product expressions, and in the calculations A_s and B_s are treated as if they were ordinary numbers and their order can be corrected afterwards.

From the formula (5) in the reference (9), we get

$$\exp[A+B] = I_0 + I_1 + I_2 + \cdots,$$

$$I_0 = \exp[\int_0^1 A_s ds] = \exp[A];$$

$$I_1 = \exp[\int_0^1 A_s ds]\int_0^1 B_{s'}\,ds',$$

$$I_2 = \frac{1}{2!}\left(\exp[\int_0^1 A_s ds]\right)\int_0^1 B_{s'}\,ds'\int_0^1 B_{s''}\,ds'',$$

$$I_3 = \frac{1}{3!}\left(\exp[\int_0^1 A_s ds]\right)\int_0^1 B_{s'}\,ds'\int_0^1 B_{s''}\,ds''\int_0^1 B_{s'''}\,ds''',\tag{A$\cdot$2}$$

$$\cdots\cdots\cdots\cdots$$

In I_2, I_3, $\cdots$ we can change the domains of the multiple integrations and carry out them as follows:

$$I_2 = \exp[\int_0^1 A_s ds]\int_0^1 B_{s'}\,ds'\int_0^{s'} B_{s''}\,ds''$$

$$= \int_0^1 ds'\,\exp[(1-s')A_4]\,B_3\int_0^{s'}\exp[(s'-s'')A_2]\,B_1\,\exp[s''A_0]\,ds''$$

$$= \exp[A_1]\int_0^1 \exp[s'(A_2 - A_1)]\frac{\exp[s'(A_0 - A_2)] - 1}{A_0 - A_2}\,ds'\,B_3 \cdot B_1$$

$$= \frac{1}{A_2 - A_0}[E(4,2) - E(4,0)]B_3 \cdot B_1,$$

$$E(x,y) = \frac{\exp[A_x] - \exp[A_y]}{A_x - A_y}.\tag{A$\cdot$3}$$

In the similar way, we obtain

$$I_3 = \exp[\int_0^1 A_s ds]\int_0^1 B_{s'}\,ds'\int_0^{s'} B_{s''}\,ds''\int_0^{s''} B_{s'''}\,ds'''$$

$$= \frac{1}{A_0 - A_2}\left[\frac{1}{A_0 - A_4}[E(0,6) - E(4,6)] - \frac{1}{A_2 - A_4}[E(2,6) - E(4,6)]\right]B_5B_3B_1,$$

(A·4)

and

$$I_4 = \frac{1}{A_0 - A_2}\left[\frac{1}{A_0 - A_4}\left[\frac{1}{A_0 - A_6}[E(0,8) - E(6,8)] - \frac{1}{A_4 - A_6}[E(4,8) - E(6,8)]\right]\right.$$

$$\left. - \frac{1}{A_2 - A_4}\left[\frac{1}{A_2 - A_6}[E(2,8) - E(6,8)] - \frac{1}{A_4 - A_6}[E(4,8) - E(6,8)]\right]\right]B_7B_5B_3B_1.$$

(A·5)

We can proceed as far as we want.

Now we take the representation in which the operator A has the diagonal form and rearrange the above results. We have

$$(I_0)_{\nu\nu'} = e^{A_\nu}\delta_{\nu\nu'},$$

$$(I_1)_{\nu\nu'} = \left(\frac{e^{A_\nu}}{(A_\nu - A_{\nu'})} + \frac{e^{A_{\nu'}}}{(A_{\nu'} - A_\nu)}\right)B_{\nu\nu'},$$

$$(I_2)_{\nu\nu'} = \sum_\sigma\left(\frac{e^{A_\nu}}{(A_\nu - A_\sigma)(A_\nu - A_{\nu'})} + \frac{e^{A_\sigma}}{(A_\sigma - A_\nu)(A_\sigma - A_{\nu'})}\right.$$

$$\left. + \frac{e^{A_\sigma}}{(A_{\nu'} - A_\nu)(A_{\nu'} - A_\sigma)}\right)B_{\nu\sigma}B_{\sigma\nu'},$$

(A·6)

$$(I_3)_{\nu\nu'} = \sum_\rho\sum_\sigma\left(\frac{e^{A_\nu}}{(A_\nu - A_\rho)(A_\nu - A_\sigma)(A_\nu - A_{\nu'})} + \frac{e^{A_\rho}}{(A_\rho - A_\nu)(A_\rho - A_\sigma)(A_\rho - A_{\nu'})}\right.$$

$$\left. + \frac{e^{A_\sigma}}{(A_\sigma - A_\nu)(A_\sigma - A_\rho)(A_\sigma - A_{\nu'})} + \frac{e^{A_{\nu'}}}{(A_{\nu'} - A_\nu)(A_{\nu'} - A_\rho)(A_{\nu'} - A_\sigma)}\right)B_{\nu\rho}B_{\rho\sigma}B_{\sigma\nu'}.$$

Proceeding in this way, we get

$$(I_n)_{\nu_0\nu_n} = \sum\sum\cdots\sum E_n(\nu_0, \nu_1, \cdots\nu_n)B_{\nu_0\nu_1}B_{\nu_1\nu_2}\cdots B_{\nu_{n-1}\nu_n}$$

$$E_n(\nu_0, \nu_1, \cdots, \nu_n) = \sum_{i=0}^{n}\frac{e^{A_i}}{\prod_{k(\neq i)}(A_i - A_k)}.$$

(A·7)

This turns out to be the special case of the Schafloth's results.[10] And a slight extension of his formula gives,

$$\lim E(\nu_0, \nu_1, \cdots, \nu_n) = \frac{1}{\lambda!\mu!}\frac{\partial^{\lambda+\mu}}{\partial A_0^\lambda\partial A_1^\mu}E_{n-\mu-\lambda-1}(\nu_0, \nu_1, \nu_{m1}, \cdots \nu_{mp}),$$

$$\nu_{k1}\rightarrow\nu_0 \qquad \nu_{l1}\rightarrow\nu_1$$
$$\vdots \qquad\qquad \vdots \qquad\qquad (p = n-\lambda-\mu-1)$$
$$\nu_{k\lambda}\rightarrow\nu_0 \qquad \nu_{l\mu}\rightarrow\nu_1$$

(A·8)

where the numbering of ν_k is immaterial.

Then we get from (A·6)

$$(I_0)_{\nu\nu} = \exp\left[A_\nu\right](I_1)_{\nu\nu} = [A_\nu]\,B_{\nu\nu}\,, \tag{A·9}$$

$$(I_2)_{\nu\nu} = \frac{1}{2!}\frac{\partial^2}{\partial A_\nu^2}\exp\left[A_\nu\right]B_{\nu\nu}^2 + \sum_{\sigma(\neq\nu)}\frac{\partial}{\partial A_\nu}E_1(\nu,\sigma)B_{\nu\sigma}B_{\sigma\nu}$$

$$= \frac{1}{2!}\exp\left[A_\nu\right]B_{\nu\nu}^2 + \sum_{\sigma(\neq\nu)}\left\{\frac{\exp\left[A_\nu\right]}{(A_\nu-A_\sigma)} + \frac{\exp\left[A_\sigma\right]-\exp\left[A_\nu\right]}{(A_\sigma-A_\nu)^2}\right\}B_{\nu\sigma}B_{\sigma\nu}, \tag{A·11}$$

$$(I_3)_{\nu\nu} = \frac{1}{3!}\exp\left[A_\nu\right]B_{\nu\nu}^3 + \sum_{\sigma(\neq\nu)}\left\{\frac{1}{2!}\frac{\partial^2}{\partial A_\nu^2}E_1(\nu,\sigma)B_{\nu\sigma}B_{\sigma\nu}B_{\nu\nu}\times 2\right.$$

$$\left. + \frac{\partial^2}{\partial A_\nu\partial A_\sigma}E_1(\nu,\sigma)B_{\nu\sigma}B_{\sigma\sigma}B_{\sigma\nu}\right\} + \sum_{\substack{\rho\neq\sigma\\(\neq\nu)}}\frac{\partial}{\partial A_\nu}E_2(\nu,\rho,\sigma)B_{\nu\rho}B_{\rho\sigma}B_{\sigma\nu} \tag{A·11}$$

$$= \frac{1}{3!}\exp\left[A_\nu\right]B_{\nu\nu}^3 + \exp\left[A_\nu\right]B_{\nu\nu}\sum_{\sigma(\neq\nu)}\frac{B_{\nu\sigma}B_{\sigma\nu}}{(A_\nu-A_\sigma)} +$$

$$+ \sum_{\substack{\rho\neq\sigma\\(\neq\nu)}}\left\{\frac{\exp\left[A_\nu\right]}{(A_\nu-A_\rho)(A_\nu-A_\sigma)} + \frac{E(\rho,\nu)}{(A_\rho-A_\sigma)(A_\rho-A_\nu)} + \frac{E(\sigma,\nu)}{(A_\sigma-A_\rho)(A_\sigma-A_\nu)}\right\}$$

$$\times B_{\nu\rho}B_{\rho\sigma}B_{\sigma\nu} +$$

$$+ 2\sum_{\sigma(\neq\nu)}\frac{\exp\left[A_\sigma\right]-\exp\left[A_\nu\right]}{(A_\sigma-A_\nu)^3}(B_{\nu\nu}B_{\nu\sigma}B_{\sigma\nu} - B_{\nu\sigma}B_{\sigma\sigma}B_{\sigma\nu}) +$$

$$- \sum_{\sigma(\neq\nu)}\frac{1}{(A_\sigma-A_\nu)^2}(\exp\left[A_\nu\right]B_{\nu\nu}B_{\nu\sigma}B_{\sigma\nu} - \exp\left[A_\sigma\right]B_{\nu\sigma}B_{\sigma\sigma}B_{\sigma\nu}) +$$

$$- \sum_{\sigma(\neq\nu)}\frac{\exp\left[A_\nu\right]}{(A_\sigma-A_\nu)^2}(B_{\nu\nu}B_{\nu\sigma}B_{\sigma\nu} - B_{\nu\sigma}B_{\sigma\sigma}B_{\sigma\nu}).$$

In these expressions there are many terms which vanishes by the trace calculation as will be seen easily.

We can go further in the similar way, though it becomes very troublesome. But we can see that $(I_n)_{\nu\nu}$ contains the term

$$(1/n!)\exp\left[A_\nu\right]B_{\nu\nu}^n. \tag{A·12}$$

If we set $A = -\beta H^{(0)}$ and $B = -\beta U$, we obtain the formula which are used in (1·11).

References

1) J. G. Kirkwood, Phys. Rev. **44** (1933), 31.
2) C. E. Uhlenbeck and B. Kahn, Physica **5** (1938), 399.
3) C. E. Uhlenbeck and E. Beth, Physica **3** (1936), 729 ; **4** (1937), 915.
4) J. E Mayer and W. Band, J. Chem. Phys. **15** (1947), 141.
5) H. S. Green, J. Chem. Phys. **19** (1951), 955 ; **20** (1952), 1247.
6) M. L. Goldberger and E. N. Adams II, J. Chem. Phys. **20** (1952), 240.
7) V. Fock, Zeits. f. Physik **75** (1932), 622.
8) I. Schiff, *Quantum Mechanics* (1949) page 340.
9) R. P. Feynman, Phys. Rev. **84** (1950), 534.
10) M. R. Shafloth, Physica **19** (1952), 645.
11) Proceedings of the International Conference on Theoretical Physics 1953. (in-press)

A Method in Quantum Staitistical Mechanics, II

—— The Degenerate Fermi-Dirac Assembly ——

Hiroshi ICHIMURA

Physics Department, Tokyo Institute of Technology,

Oh-Okayama, Meguroku, Tokyo

Using the method which has been developed by the author, we calculate the effect of the inter-particle interaction of the thermodynamic potential per particle of the highly degenerated Fermi-Dirac gas. We show that under certain assumptions the specific heat of the system of conduction electrons in a metal maintain its linear temperature dependence in the presence of the interaction.

§ 1. An assembly of interacting Fermi-Dirac particles

As an application of the formula which we have derived in Part I.[1], we want to treat the highly degenerated Fermi gas in which the parameter λ may be considered to be very large. The interaction potential between particles is assumed to have the following cut-offed Coulomb type:

$$JG(r) = J\frac{e^{-\alpha r}}{r} \quad (J > 0, \alpha > 0). \tag{1·1}$$

Although there may not be any actual physical system which possesses this interaction as an elementary interaction, we can take this Fermi gas for a model for the system of conduction electrons in the metal. As a result of our treatment we can reconfirm rather clearly the conclusion given by Wohlfarth[2] by the same model.

The formula (1·13) in Part I (we use the notation such as I—(1·13) hereafter) for the grand partition function Z of our system becomes in this case as

$$Z = \prod_r (1 + \lambda e^{-\beta \varepsilon r}) \exp\left[\sum J^i C_i\right] \tag{1·2}$$

and we get

$$\bar{N} = \sum \frac{\lambda e^{-\beta \varepsilon r}}{1 + \lambda e^{-\beta \varepsilon r}} + \sum J^i \lambda \frac{\partial}{\partial \lambda} C_i . \tag{1·3}$$

If the interaction vanishes ($J = 0$), we get from (1·3) the usual formula for the thermodynamic potential per particle $\mu (= \kappa T \log \lambda)$, that is

$$\mu = \mu_0\left(1 - \frac{\pi^2}{12}\left(\frac{\kappa T}{\mu_0}\right)^2 - \frac{\pi^4}{80}\left(\frac{\kappa T}{\mu_0}\right)^4 + \cdots\right), \quad \mu_0 = \frac{h^2}{8m}\left(\frac{6N}{\pi V}\right)^{2/3}, \tag{1·4}$$

where g means the spin weight and we take $g=2$ for our case. μ_0 is the energy of the Fermi surface at O°K for the ideal Fermi gas. In the following, we calculate the corrections for this formla which appear on account of the presence of the terms $\sum J^2\lambda\dfrac{\partial}{\partial\lambda}C_i$.

§ 2. Evaluation of C_1 and C_2

(1) C_1: From I-(1·1), I-(1·2) and I-(3·3) we get

$$JC_1=-\frac{4^2\pi J}{V\alpha^2}\beta\sum_{j<k}\sum\left(1-\frac{\alpha^2}{2(\alpha^2+k_{jk}^2)}\right)f_jf_k,\ f_j=\frac{\lambda e^{-\beta\varepsilon_j}}{1+\lambda e^{-\beta\varepsilon_j}},\qquad(2\cdot1)$$

where we have used the results

$$JG_{jk}=\frac{4^2\pi J}{V\alpha^2}\left(1-\frac{\alpha^2}{2(\alpha^2+k_{jk}^2)}\right)\qquad(2\cdot2)$$

which is well known for the potential (1·1).

If we consier the case of the usual metal and assume that α is of the order $10^8\mathrm{cm}^{-1}$ (this means the cut-off distance is the order of $1°A$), or more directly, if we assume $\dfrac{\hbar^2\alpha^2}{2m}>\mu$, we can calculate (2·1) as

$$JC_1=-N\frac{8\pi J N}{V\alpha^2}\beta\left(\frac{\mu}{\mu_0}\right)^3\left\{\frac{1}{2}\left(1+\frac{\pi^2}{4}\left(\frac{\kappa T}{\mu}\right)^2-\frac{3\pi^4}{80}\left(\frac{\kappa T}{\mu}\right)^4+\cdots\right)\right.$$

$$\left.+\frac{3}{5}\frac{2\mu m}{\hbar^2\alpha^2}\left(1+\frac{3\pi^2}{4}\left(\frac{\kappa T}{\mu}\right)^2+\cdots\right)+O\left(\left(\frac{2m\mu}{\hbar^2\alpha^2}\right)\right)\right\},$$

where we have replaced the summation $\sum_j$ with the integral

$$\frac{V}{(2\pi)^3}\int d\boldsymbol{k}$$

(spin summations are already carried out), and used the well known approximation method for the Fermi distribution function. (Appendix I)

Considering the complications involved in the calculation of C_2 and others, however, we content ourselves with the following approximate treatment. But it will be sufficient for the purpose of the order estimation of the correction terms in μ. We replace all

$$1-\frac{\alpha^2}{2(\alpha^2+k_{jk}^2)}$$

with a constant factor $\gamma(<1)$, so the summation can easily be performed,

$$JC_1=-\frac{1}{2}\beta\gamma\frac{4^2\pi J}{V\alpha^2}\left(\sum_j\frac{\lambda e^{-\beta\varepsilon_j}}{1+\lambda e^{-\beta\varepsilon_j}}\right)^2$$

$$= -2N\gamma\theta\beta\left(\frac{\mu}{\mu_0}\right)^3\left(1 + \frac{\pi^2}{4}\left(\frac{\kappa T}{\mu}\right)^2 + \frac{3\pi^4}{80}\left(\frac{\kappa T}{\mu}\right)^4 + \cdots\right), \tag{2.3}*$$

$$\theta = \frac{4\pi J/N}{V\alpha^2}.$$

From this we obtain

$$J\lambda\frac{\partial}{\partial\lambda}C_1 = \kappa TJ\frac{\partial}{\partial\mu}C_1$$

$$= -3N\gamma\frac{\theta}{\mu_0}\left(\frac{\mu}{\mu_0}\right)^2\left(1 + \frac{\pi^2}{12}\left(\frac{\kappa T}{\mu^2}\right)^2 - \frac{3\pi^4}{80}\left(\frac{\kappa T}{\mu}\right)^4 + \cdots\right). \tag{2.4}$$

(2) C_2; In the similar way, we get from I-(2.5), I-(2.15) and I-(3.3)

$$C_2 = D_2 + \langle N_2\rangle,$$

$$D_2 = \frac{\beta^2J^2}{2}\sum_{j\neq k\neq l}\sum\sum\frac{G_{jk}G_{jl}}{(1+\lambda e^{-\beta\varepsilon_j})}f_jf_kf_l(1-f_j)$$

$$= \frac{\beta^2\gamma^2}{2}\left(\frac{4^2\pi J}{V\alpha^2}\right)\left(\sum_j\frac{1}{(1+\lambda e^{-\beta\varepsilon_j})(1+\lambda^{-1}e^{\beta\varepsilon_j})}\right)\left(\frac{\lambda e^{-\beta\varepsilon_k}}{1+\lambda e^{-\beta\varepsilon_k}}\right)^2 \tag{2.5}$$

$$= \frac{3}{8}N\theta^2\gamma^2\frac{\beta}{\mu_0}\left(\frac{\mu}{\mu_0}\right)^{7/2}\left(1 + \frac{\pi_2}{6}\left(\frac{\kappa T}{\mu}\right)^2 + \frac{67\pi^4}{1920}\left(\frac{\kappa T}{\mu}\right)^4 + \cdots\right)$$

and

$$\lambda\frac{\partial}{\partial\lambda}D_2 = \frac{21}{16}\left(\frac{\theta}{\mu_0}\gamma\right)^2\left(\frac{\mu}{\mu^2}\right)^{5/2}\left(1 + \frac{\pi^2}{4}\left(\frac{\kappa T}{\mu}\right)^2 - \frac{67\pi^4}{13440}\left(\frac{\kappa T}{\mu}\right)^4 + \cdots\right) \tag{2.6}$$

and from I-(2.14), it becomes

$$J^2\langle N_2\rangle = J^2\beta^2 S_{pq,rs}\frac{(G_{pq,rs})^2}{\beta(\varepsilon_p+\varepsilon_q-\varepsilon_r-\varepsilon_s)}\frac{1}{(1+\lambda e^{-\beta\varepsilon_p})(1+\lambda e^{-\beta\varepsilon_q})}f_rf_s$$

$$= \frac{J^2\beta^2}{2\lambda^2}S_{pq,rs}\frac{(G_{pq,rs})^2}{\beta(\varepsilon_p+\varepsilon_q-(\varepsilon_r+\varepsilon_s))}f_rf_sf_pf_q, \tag{2.7}$$

where $(G_{pq,rs})$ takes the following form, as can be evaluated from the definition in I-(1.1):

$$(G_{pq,rq}) = \frac{4\pi}{V}\left(\frac{\theta_{pq,rs}}{\alpha^2+k_{sq}^2} - \frac{\theta_{pq,sr}}{\alpha^2+k_{rq}^2}\right). \tag{2.8}$$

Now we want to show $J^2\langle N_2\rangle$ is negligible in comparison with D_2. Replacing $(G_{pq,rs})$ with the larger quantity $4\pi/V\alpha^2$, we obtain

$$J^2\langle N_2\rangle \lesssim \frac{\beta^2}{2\lambda^2}\left(\frac{4\pi J}{V\alpha^2}\right)^2 S_{pq,rs}B_{pq,rs}f_rf_sf_pf_q,$$

$$B_{pq,rs} = \frac{e^{\beta\varepsilon_{pq}} - e^{\varepsilon\beta rs}}{\beta(\varepsilon_{pq}-\varepsilon_{rs})}, \quad \varepsilon^{xy} = \varepsilon_x + \varepsilon_y. \tag{2.9}$$

---

* We use N instead of $\mathscr{N}$ hereafter.

The exact evaluation of $(2 \cdot 9)$ seems to be very difficult.   But we can show by an approximate treatment that the following inequality holds : (Appendix II)

$$J^2 \langle N_2 \rangle < N \frac{3^2}{2^5} \Big(\frac{\theta}{\mu_0} \Big)^2 \Big(\frac{\mu}{\mu_0} \Big)^{5/2}. \tag{2·10}$$

(3)   Other terms :   By the similar reasoning, we know

$$J^3 D_{\cdot} \sim \Big(\frac{\theta}{\mu_0} \Big)^3,$$

$$J^3 \langle F_3 \rangle - \langle D \rangle \langle N_2 \rangle \gtrsim \sim J^2 \langle N_2 \rangle \frac{\theta}{\mu_0} \tag{2·11}$$

and for $J^4 C_4$, etc. we find corresponding relations.

§ 3.   The thermodynamic potential per particle : μ

With the above estimation for $\sum J^i C_i$, the equation can be evaluated as

$$N = N \Big(\frac{\mu}{\mu_0} \Big)^{3/2} \Big(1 + \frac{\pi^2}{8} z + \frac{7\pi^4}{640} z^2 + \cdots \Big)$$

$$- N \frac{3}{2} \gamma \frac{\theta}{\mu_0} \Big(\frac{\mu}{\mu_0} \Big)^2 \Big(1 + \frac{\pi^2}{12} z - \frac{3\pi^4}{80} z^2 + \cdots \Big)$$

$$+ O\Big(N \Big(\frac{\theta}{\mu_0} \Big)^2 \Big), \tag{3·1}$$

$$z = \Big(\frac{\kappa T}{\mu} \Big)^2.$$

From this, we know μ approximately to the second order in z and to the first order in θ/μ_0.   After a lengthy but straight-forward calculation, we obtain

$$\mu = \mu_0 \Big\{ 1 + 2 \Big(\frac{\theta}{\mu_0} \Big) + O \Big(\Big(\frac{\theta}{\mu_0} \Big)^2 \Big)$$

$$- \frac{\pi^2}{12} \Big(\frac{\kappa T}{\mu_0} \Big)^2 \Big(1 + O \Big(\Big(\frac{\theta}{\mu_0} \Big)^2 \Big) \Big) \tag{3·2}$$

$$- \frac{\pi^4}{80} \Big(\frac{\kappa T}{\mu_0} \Big)^4 \Big(1 - \frac{5}{72} \Big(\frac{\theta}{\mu_0} \Big) + O \Big(\Big(\frac{\theta}{\mu_0} \Big)^2 \Big) \Big) \Big\}.$$

From this μ, we can calculate the free energy $\phi(V. T)$ by

$$\phi(V, T) = \frac{1}{N} \int_0^N \mu(N) \, dN. \tag{3·3}$$

Considering the N dependence of μ_0 and θ, we get

$$\phi(V, T) = \mu_0 \left\{ \frac{3}{5} + \left(\frac{\theta}{\mu_0} \right) + O\left(\left(\frac{\theta}{\mu_0} \right)^2 \right) \right.$$

$$- \frac{\pi^2}{4} \left(\frac{\kappa T}{\mu_0} \right)^2 \left(1 + O\left(\left(\frac{\theta}{\mu_0} \right)^2 \right) \right)$$

$$\left. + \frac{\pi^4}{80} \left(\frac{\kappa T}{\mu_0} \right)^4 \left(1 + \frac{5}{48} \left(\frac{\theta}{\mu_0} \right) + O\left(\left(\frac{\theta}{\mu_0} \right)^2 \right) \right) \right). \tag{3.4}$$

If the quantity θ/μ_0 is sufficiently small, we can conclude that our procedures converge and the linearity of the temperature dependence of the specific heat of the degenerate Fermi gas is not disturbed by the presence of the inter-particle interaction $(1 \cdot 1)$.

If we set $\mu_0 \sim 10^{11}$ erg, $N = \nu \times 6.02 \times 10^{23}$ (ν; the number of conduction electron per atom), $V \sim 10$cc, as is the case with ordinary metals and assume

$$J = (\delta e)^2 \quad (\delta \text{ is screening constant, } e = 4.80 \times 10^{-10} \text{ esu}),$$

$$a = 10^8 \text{ cm}^{-1} \quad (\text{this means the cut-off length is } 10^8 \text{ cm}),$$

it becomes

$$\frac{\theta}{\mu^0} \sim \nu \delta^2. \tag{3.5}$$

So the assumption $\delta \sim 0.1$ will be sufficient for our discussions. As a matter of fact, it is too hasty to conclude merely from our reasoning that the potential of type $(1 \cdot 1)$ with these constants is really acting between the conduction electrons. But, considering the situation that the electrons in the metal are moving in the sea of the positive charges, and afterall, the extreme difficulties of treating the long range Coulomb interaction, the present model will be helpfull for the discussions concerning the effects of the inter-electronic interaction on the thermal behaviour of the conduction electrons,

Appendix I. Calculation of the equations $(2 \cdot 1)$ and $(2 \cdot 5)$

In $(2 \cdot 1)$ we must evaluate the integral

$$I = a^2 \iint \frac{1}{a^2 + q^2} \frac{d\mathbf{k}}{1 + \exp[\beta(sk^2 - \mu)]} \frac{d\mathbf{k}'}{1 + \exp[\beta(sk'^2 - \mu)]} \tag{A1.1}$$

where $s = \hbar^2/2m$ and $q = |\mathbf{k}' - \mathbf{k}|$.

Introducing the variables as in Fig. 1., we get

$$I = a^2 \int d\mathbf{k}' \int_0^\infty \int_0^\pi \int_0^{2\pi} k^2 \sin \theta \, dk \, d\theta \, d\varphi$$

$$\times \frac{1}{a^2 + k^2 + k'^2 - 2kk' \cos \theta} g(k^2) g(k'^2)$$

$$g(k^2) = \frac{1}{1 + \exp[\beta(sk^2 - \mu)]} \tag{A1.2}$$

(k', θ', φ')

(k, θ, φ)

θ
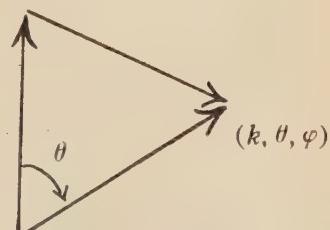
Fig. 1

We perform the integral over θ, φ, θ', φ', and change the variables k to ε and k' to ε' by

$$ck^2 = \varepsilon, \quad sk'^2 = \varepsilon'.$$

So it becomes

$$I = \alpha^2 \frac{(2\pi)^2}{(2s)^2} \int_0^\infty \int_0^\infty \log \frac{(\alpha^2 + \varepsilon/s + \varepsilon'/s + 2/s\sqrt{\varepsilon}\,\sqrt{\varepsilon'}\,)}{(\alpha^2 + \varepsilon/s + \varepsilon'/s - 2/s\sqrt{\varepsilon}\,\sqrt{\varepsilon'}\,)} g(\varepsilon)g(\varepsilon')d\varepsilon d\varepsilon'. \quad (A1 \cdot 3)$$

Now we use the well known expansion technique

$$\int_0^\infty A(\varepsilon)f(\varepsilon)d\varepsilon = \int_0^\mu A(\varepsilon)d\varepsilon + \frac{\pi^2}{6}(\kappa T)^2\left(\frac{dA}{d\varepsilon}\right)_\mu + \frac{7\pi^4}{360}(\kappa T)^4\left(\frac{d^3A}{d\varepsilon^3}\right)_\mu + \cdots$$

and expand the logarithm with the assumption $\dfrac{\hbar^2}{2m}\alpha^2 > \mu.$

So we obtain

$$I = \frac{2\pi}{s^3}\left(\frac{2}{3}\right)^2\mu^3\left[\left(1 + \frac{\pi^2}{8}\left(\frac{\kappa T}{\mu}\right)^2 + \frac{7\pi^4}{640}\left(\frac{\kappa T}{\mu}\right)^4 + \cdots\right)\right.$$

$$-\frac{3}{5}\frac{2\mu m}{\hbar^2\alpha^2}\left(1 + \frac{3}{4}\pi^2\left(\frac{\kappa T}{\mu}\right)^2 + \frac{43}{480}\pi^4\left(\frac{\kappa T}{\mu}\right)^4 + \cdots\right)$$

$$\left. + O\left(\left(\frac{2m\mu}{\hbar^2\alpha^2}\right)^2\right)\right]. \quad\quad\quad (A1 \cdot 4)$$

It can be seen that that the first term of this equation is one half of the another term in (2·1). Then our result is obtained.

In (2·5), it will be sufficient if we give the value of

$$K = \sum_j \frac{1}{(1 + \lambda c^{-\beta\varepsilon_j})(1 + \lambda^{-1}c^{\beta\varepsilon_j})}.$$

This can be evaluated by the same method and we obtain

$$K = \frac{3}{4}N\frac{\beta}{\mu_0}\left(\frac{\mu}{\mu_0}\right)^{1/2}\left(1 - \frac{\pi^2}{12}\left(\frac{\kappa T}{\mu}\right)^2 - \frac{7\pi^4}{384}\left(\frac{\kappa T}{\mu}\right)^4 + \cdots\right).$$

Appendix II. Proof of (2·10)

We replace the summation $S_{pq,rs}$ on the right hand side of the equation (2·9) by the integral (spin summation yield the factor 4), then we have

$$J^2\langle N_2\rangle < M,$$

$$M = \frac{\varepsilon^2}{2\lambda^2}\left(\frac{4\pi J}{Va^2}\right)^2\frac{4}{4}\left(\frac{3}{4}N\frac{1}{\mu_0}\right)^{3/2} \quad\quad (A2 \cdot 1)$$

$$\times \int\int\int B_{pq,rs}f_rf_sf_p\,t_q\sqrt{\varepsilon_q}d\varepsilon_p\sqrt{\varepsilon_q}d\varepsilon_q\sqrt{\varepsilon_r}d\varepsilon_r,$$

where

$$\varepsilon_s = \frac{\hbar^2}{2m}(k_s)^2 = \frac{\hbar^2}{2m}(k_p + k_q - k_r)^2,$$

Making use of the well known property of the functions f_p, etc and a integral represen- tation for $B_{pq,rs}$, we get approximately

$$M = \frac{N}{2\lambda^2}\left(\frac{4\pi JN}{V\omega^2}\beta\right)^2\left(\frac{3}{4}\mu_0^{-3/2}\right)^3\int_0^1 d\rho J(\rho),$$

$$J(\rho) = \int\int_0^\mu\int e^{\beta(\varepsilon_p + \varepsilon_p)}e^{\rho\beta(\varepsilon_r + \varepsilon_s - \varepsilon_p - \varepsilon_q)}\sqrt{\varepsilon_p}\, d\varepsilon_p\sqrt{\varepsilon_q}\, d\varepsilon_q\sqrt{\varepsilon_r}\, d\varepsilon_r. \qquad (A2 \cdot 2)$$

$J(\rho)$ has the following properties:

$$J(0) = \int\int_0^\mu\int e^{\beta(\varepsilon_p + \varepsilon_q)}\sqrt{\varepsilon_p}\, d\varepsilon_p\sqrt{\varepsilon_q}\, d\varepsilon_q\sqrt{\varepsilon_r}\, d\varepsilon_r$$

$$= \int\int_0^\mu\int e^{\beta(\varepsilon_p + \varepsilon_q)}\sqrt{\varepsilon_p}\, d\varepsilon_p\sqrt{\varepsilon_q}\, d\varepsilon_q\sqrt{\varepsilon_r}\, d\varepsilon_r = J(1), \qquad (A2 \cdot 3)$$

where the relations between variables as in $(A2 \cdot 1)$ holds. And the relation

$$\frac{d^2J}{d\rho^2} = \int\int_0^\mu\int\beta^2(\varepsilon_{rs} - \varepsilon_{pq})^2 e^{\beta(\varepsilon_p + \varepsilon_q)}e^{\rho\beta(\varepsilon_{rs} - \varepsilon_{pq})}\sqrt{\varepsilon_p}\, d\varepsilon_p\sqrt{\varepsilon_q}\, d\varepsilon_q\sqrt{\varepsilon_r}\, d\varepsilon_r \geq 0 \quad (A2 \cdot 4)$$

holds all over $(0, 1)$.

Then we can conclude

$$J(\rho) \leq J(0), \quad 0 \leq \rho \leq 1$$

and the following inequality is satisfied

$$M \leq \frac{N}{2\lambda^2}\left(\frac{4\pi JN}{V\omega^2}\beta\right)^2\left(\frac{3}{4}\mu_0^{-3/2}\right)^3 J(0). \qquad (A2 \cdot 5)$$

We can evaluate $J(0)$ easily as

$$J(0) = \int\int_0^\mu\int e^{\beta(\varepsilon_p + \varepsilon_q)}\sqrt{\varepsilon_p}\, d\varepsilon_p\sqrt{\varepsilon_p}\, d\varepsilon_p\sqrt{\varepsilon_r}\, d\varepsilon_r = \left(\int_0^\mu e^{\beta\varepsilon}\sqrt{\varepsilon}\, d\varepsilon\right)^2\frac{2}{3}\mu^{3/2}, \qquad (A2 \cdot 6)$$

where

$$\int_0^\mu e^{\beta\varepsilon}\sqrt{\varepsilon}\, d\varepsilon = \left(\frac{1}{\beta\mu}e^{\beta\mu} - \frac{1}{\beta\mu^{3/2}}\int_0^{\sqrt{\mu}}e^{\beta x^2}dx\right)\mu^{3/2}.$$

Then it becomes

$$J(0) < \frac{2}{3}e^{2\beta\mu}(\mu^{3/2})^3\frac{1}{\beta\mu}.$$

So we get finally,

$$M \leq N(\theta/\mu_0)^2(3^2/2)^5((\mu/\mu_0)^{5/2}. \qquad (A2 \cdot 7)$$

References

1) H. Ichimura, Prog. Theor. Phys. **11** (1954), 384.
2) E. P. Wohlfarth, Phil. Mag. **91** (1950), 534.

The Problem of Many Bodies and the Superposition of Spherically Symmetric Space-times in General Relativity

Hyôitirô TAKENO

Research Institute for Theoretical Physics, Hiroshima University

(Takehara-machi, Hiroshima-ken)

(Received October 20, 1953)

We put forward a possibility of solving the problem of many bodies in general relativity by introducing the concept of the superposition of the space-times. And we construct an invariant theory concerning the superposition of the static spherically symmetric space-times having the centers of symmetry in common. Namely, we give the definitions of two operations: (i) multiplication of a space-time by a constant and (ii) summation of two space-times, invariant-theoretically, and study their consequences. We give some examples using the space-times frequently treated in general relativity. It is also shown that we can introduce invariantly various operations concerning these space-times beside the above two. Lastly another important rôle which this theory of the superposition may play in general relativity is considered.

§ 1. Introduction

One body problem in the general theory of relativity was completely solved in an exact form by Schwarzschild's exterior solution of the Einstein's field equation

$$K_{ij} - 1/2 \cdot K g_{ij} = 0, \qquad (i, j = 1, \cdots, 4),^* \qquad (1 \cdot 1)$$

where K_{ij} and K are Ricci tensor and the scalar curvature of the space-time respectively. On the other hand, the problem of many bodies in the theory has not yet been solved exactly although many efforts have been made concerning this problem. Most of the investigations were attempted by using approximate methods and only few researchers tried to obtain the exact solution.**

Now we shall confine ourselves to the problem of two bodies. The most natural method of solving the problem exactly may be to search for the axially symmetric solution of (1.1).*** We can also take the special method of using the Eddington's wave tensor calculus and that of using a Riemannian space of higher dimension following after the one used in the same problem in the relativistic quantum mechanics.**** But we have not yet

---

* In this paper the indices $i, j, k, l, \cdots$ take the values $1, \cdots, 4$.

** To solve the problem approximately we have the following two methods: The first is the somewhat 'old' one initiated by Einstein and developed by de Sitter, Droste, Levi-Civita, etc., and the second is the comparatively new one which lays stress on the field equation alone and was developed by Einstein, Infeld, Hoffmann, etc.

*** In connection with this method we have the research of Silberstein[1] concerning the axially symmetric solution of Weyl and Bach.

**** Hosokawa studied the problem of many bodies in Wave Geometry using this method.[2]

succeeded. Anyhow it is very difficult to solve the problem exactly in general relativity.

In this paper we intend to advocate the possibility of solving the problem from a new point of view by introducing the concept of the ' superposition of the space-times '. In the following we shall explain this somewhat in detail with respect to the problem of two bodies.

Let M_1 and M_2 be two massive particles whose masses are m_1 and m_2 respectively. Then the gravitational field due to M_1 alone or M_2 alone is completely expressible by using the Schwarzschild's exterior solution. We shall denote these two space-times by R_1 and R_2 respectively. Of course both R_1 and R_2 are static and spherically symmetric and M_1 and M_2 are located in the respective centers of symmetry. From this fact we assert that there is a possibility of solving the problem of two bodies by superposing these R_1 and R_2 suitably. Let R_3 be the resulting space-time obtained by this procedure. It is required naturally that the superposition must satisfy the following two conditions. The first is that when the positions of M_1 and M_2 coincide with each other the resulting R_3 must be the same Schwarzschild's space-time corresponding to the mass $(m_1 + m_2)$. As will be seen later our theory developed in this paper satisfies this condition completely. When the positions of M_1 and M_2 do not coincide they will make some relative motion and the space-time R_3 which is to be determined by R_1 and R_2 will be neither spherically symmetric nor static in general, and further this R_3 must represent in a suitable form the interaction between the gravitational fields due to M_1 and M_2. This is the second condition. By our conjecture such an interaction is to be expressed by some relations among the curvature tensors of the three space-times R_1, R_2 and R_3. Our assertion is that there is a possibility of solving the problem of many bodies by introducing the concept of the superposition of the space-times satisfying such requirements.

It should be noticed, however, that when we proceed to solve the problem by the new method above stated we can not assure that the resulting space-time R_3 would satisfy the field equation (1.1). By the usual methods hitherto considered the solution of the problem is to be attained by solving the field equation directly. On the other hand, by our new way the curvature tensor of the resulting space-time will not necessarily satisfy $(1 \cdot 1)$ though each of the constituent space-times satisfies it, since the superposition will not be linear with respect to the Ricci tensor. Namely the R_3 may not satisfy (1.1) on account of the interaction between R_1 and R_2. Accordingly our new method will not necessarily bring the complete solution in the sense hitherto considered.

In the above we have made clear that the concept of the superposition of space-times is closely related with the solution of the problem of many bodies in the general theory of relativity. Nevertheless it seems to the writer that this problem of how we can obtain a space-time by superposing some space-times has never been treated as yet. In some researches concerning approximate solution of the problem of many bodies there are some results which bear some relation to this problem in dealing with the combined effect of many weak gravitational fields.[3] But no writer has ever mentioned explicitly about the superposition of space-times.

In this paper we are going to construct a theory concerning the superposition of space-

times, but the curvature tensor of a space-time is a complicated non-linear function of the fundamental tensor g_{ij} of the space-time and accordingly it is not easy in general to obtain some simple relations among respective fundamental tensors, Ricci tensors, energy-momentum tensors, etc. of the space-times by defining the superposition suitably. When the space-times we deal with are restricted appropriately, however, it is not impossible to construct a theory of the superposition satisfying this requirement. In fact, in this paper, we shall be able to construct a theory with respect to the static spherically symmetric space-times in the highly special case in which all space-times possess their centers of symmetry in common, as a starting point of the general theory of the superposition which will be accomplished in future. Of course what we desire most to construct is a theory for the case in which the centers of symmetry are not common. In the present stage, however, it is not easy, so we shall confine ourselves to the special case above stated. Accordingly, in this paper, we do not also deal with the problem of how to apply the theory of superposition to the problem of many bodies, leaving it for future investigation, and we only give the mathematical theory of superposition for the special case.

Lastly we have to notice that our theory is constructed from the standpoint that the mathematical aspect of general relativity is a theory of analytical invariants and accordingly main definitions given in this paper are of invariant meanings independent of coordinate systems.

First in § 2 we shall recollect the main properties of static spherically symmetric space-times which constitute the basis of our present research.[1] Mathematically speaking, in order to introduce the concept of the superposition of space-times, we have only to make clear the two operations: (i) Multiplication of a space-time by a constant and (ii) summation of two space-times. Therefore in § 3, § 4 and § 5 we shall define these two operations invariant-theoretically and shall study their consequences. In § 6, we shall give some examples of the results thus obtained by using the space-times frequently treated in general relativity. In § 7 we shall show that we can also introduce invariantly various operations concerning spherically symmetric space-times. Lastly in § 9, we shall consider another important rôle the theory of the superposition of space-times may play in general relativity.

§ 2. Static spherically symmetric space-times

In [4], the concept of the 'staticness' of a spherically symmetric space-time S was introduced as an intrinsic property of the space-time and some results concerning static S's were obtained. In this section we shall write down the main results obtained in [4] which are necessary for the present research.

An S has a set of intrinsic scalars $\overset{4}{\rho}, \overset{3}{\rho}, \overset{2}{\rho}, \overset{1}{\rho}$ and λ. ρ's are determined uniquely and λ is determined uniquely to within an m-transformation. A necessary and sufficient condition that an S be *static* is given by the tensor equations

$$\overset{4}{\rho}=\overset{4}{\rho}(\lambda), \qquad \overset{3}{\rho}=\overset{3}{\rho}(\lambda). \tag{2.1}$$

Static S's make one to one correspondence in an invariant manner with the pairs of $\overset{4}{\rho}(\lambda)$

and $\overset{3}{\rho}(\lambda)$. So we can express this correspondence by the equation

$$S = f\{\overset{4}{\rho}(\lambda),\ \overset{3}{\rho}(\lambda)\}. \tag{2.2}$$

$\overset{2}{\rho}(\lambda)$ and $\overset{1}{\rho}(\lambda)$ are given by the following relations from $\overset{4}{\rho}$ and $\overset{3}{\rho}$:

$$\overset{2}{\rho} = \lambda\overset{4}{\rho}{}', \quad \overset{1}{\rho} = 1/4 \cdot \lambda^2 XYZ^{-1} + \lambda(\overset{3}{\rho}{}' + \overset{4}{\rho}{}'), \tag{2.3}$$

where

$$\overset{a}{\rho}{}' = \frac{d}{d\lambda}\overset{a}{\rho}; \quad X = \overset{3}{\rho} - \lambda\overset{4}{\rho}{}', \quad Y = \overset{3}{\rho} + 2\overset{4}{\rho}, \quad Z = 1 + 1/2 \cdot \lambda^2\overset{4}{\rho} \quad (\neq 0). \tag{2.4}$$

Next, the line element of a static S is reducible to the form

$$ds^2 = -A(r)dr^2 - r^2(d\theta^2 + \sin^2\theta d\phi^2) + C(r)dt^2 \tag{2.5}$$

by taking the coordinate system suitably. The coordinate system in which this (2.5) holds is called *canonical.** Evidently this line element is static in the usual sense. In fact, our invariant definition of the staticness was introduced basing on this fact.

Throughout the present paper we shall deal only with all S's having a c.c.s. (abbreviation of the canonical coordinate system) in common exclusively. In general an S has one λ and it coincides with r in a c.c.s. However in some special static S's e.g., de Sitter space-time, besides this λ there may exist some λ's which do not coincide with r. (But by suitable m-transformations these λ's also coincide with r.) In the following we also do not deal with such λ's. Hence by these assumptions *we have $\lambda = r$ in the common c.c.s. for all static S's* treated in this paper.

In a c.c.s. the non-vanishing components of the curvature tensor $K_{ij}^{.lm}$ and $\overset{a}{\rho}$ of the static S satisfy the following relations:

$$\left\{ \begin{aligned} &\alpha = K_{12}^{.12} = K_{13}^{.13} = -A'/2A^2 r = 1/4 \cdot (\overset{2}{\rho} + 2\overset{4}{\rho}) = 1/2 \cdot (\overset{4}{\rho} + 1/2 \cdot r\overset{4}{\rho}{}'), \\ &\beta = K_{12}^{.24} = K_{34}^{.34} = C'/2ACr = 1/4 \cdot (\overset{3}{\rho} + 2\overset{4}{\rho}), \\ &\xi = K_{14}^{.14} = -\{-2C'' + C'(A'/A + C'/C)\}/4AC = 1/4 \cdot (\overset{1}{\rho} + \overset{2}{\rho} + \overset{3}{\rho} + 2\overset{4}{\rho}), \\ &\eta = K_{23}^{.23} = (1/A - 1)r^{-2} = 1/2 \cdot \overset{1}{\rho}, \end{aligned} \right. \tag{2.6}$$

$$\overset{1}{\rho} = 4(\xi + \eta - \alpha - \beta), \quad \overset{2}{\rho} = 4(\alpha - \eta), \quad \overset{3}{\rho} = 4(\beta - \eta), \quad \overset{4}{\rho} = 2\eta. \tag{2.7}$$

Accordingly $\overset{4}{\rho}(\lambda)$ and $A(r)$ make one to one correspondence, and $\overset{3}{\rho}(\lambda)$ and $C(r)$ also make one to one correspondence to within a constant multiplier of C. Now if we put $A = (1+\psi)^{-1}$, we have

$$\overset{4}{\rho} = 2\psi/r^2, \quad \overset{3}{\rho} = 2C'(1+\psi)/rC - 4\psi/r^2. \tag{2.8}$$

The following three kinds of static S's were studied in detail in [4] as having some remarkable properties. As will be seen later these special S's play important rôles in the present research.

(i) S_p i.e. the static S whose $X = 0$, which is equivalent to $\overset{2}{\rho} = \overset{3}{\rho}$. In this S_p we have

$$\overset{2}{\rho} = \overset{3}{\rho} = \lambda\overset{4}{\rho}{}', \quad \overset{1}{\rho} = 2\lambda\overset{4}{\rho}{}' + \lambda^2\overset{4}{\rho}{}'' = (\lambda^2\overset{4}{\rho}{}')'. \tag{2.9}$$

---

* This coordinate system is not determined uniqely. $C(r)$ is determined to within an arbitrary constant multiplier.

The line element of an S_p takes the form

$$ds^2 = -A(r)dr^2 - r^2(d\theta^2 + \sin^2\theta d\phi^2) + A^{-1}dt^2 \qquad (2\cdot10)$$

in the r.c.c.s. (abbreviation of reduced c.c.s.) and in this coordinate system $(2\cdot6)$ and $(2\cdot8)$ become

$$\begin{cases} \alpha = \beta = -A'/2A^2r = 1/2 \cdot (\overset{1}{\rho} + 1/2 \cdot r\overset{1'}{\rho}), \quad \eta = (A^{-1}-1)/r^2 = 1/2 \cdot \overset{4}{\rho}, \\ \xi = 1/2 \cdot C'' = 1/2 \cdot (A^{-1})'' = 1/4 \cdot (2\overset{1}{\rho} + 4r\overset{1'}{\rho} + r^2\overset{4''}{\rho}), \quad (\overset{1}{\rho} = \overset{1}{\rho}(r)), \end{cases} \qquad (2\cdot11)$$

and $\qquad \overset{1}{\rho} = 2\psi/r^2, \quad \overset{3}{\rho} = 2(r\psi' - 2\psi)/r^2, \quad (\psi' = d/dr \cdot \psi), \qquad (2\cdot12)$

respectively. Two well-known examples of S_p are given by de Sitter space-time and the space-time of Schwarzschild's exterior solution. Their line elements in the respective reduced canonical forms are given by

$$ds^2 = -(1-k^2r^2)^{-1}dr^2 - r^2(d\theta^2 + \sin^2\theta d\phi^2) + (1-k^2r^2)dt^2, \quad (\overset{1}{\rho} = -2k^2), \qquad (2\cdot13)$$

$$ds^2 = -(1-2m/r)^{-1}dr^2 - r^2(d\theta^2 + \sin^2\theta d\phi^2) + (1-2m/r)dt^2, (\overset{1}{\rho} = -4m/\lambda^3), (2\cdot14)$$

where k and m are constants. In the following we shall denote these space-time by $[$de S. $(\overset{1}{\rho} = -2k^2)]$ and $[$Sch. ex. $(\overset{1}{\rho} = -4m/\lambda^3)]$ respectively. Further when it is not necessary to indicate $\overset{4}{\rho}$ we shall use the simplified notations $[$de S.$]$, etc.

(ii) S_q i.e. the static S whose $Y = 0$. In this S_q we have

$$-\overset{1}{\rho} = \overset{2}{\rho} = \lambda\overset{1'}{\rho}, \quad \overset{3}{\rho} = -2\overset{1}{\rho}. \qquad (2\cdot15)$$

The line element in the reduced canonical form is given by

$$ds^2 = -A(r)dr^2 - r^2(d\theta^2 + \sin^2\theta d\phi^2) + dt^2. \qquad (2\cdot16)$$

In this r.c.c.s. $(2\cdot6)$ and $(2\cdot8)$ become

$$\alpha = -A'/2A^2r = 1/2 \cdot (\overset{1}{\rho} + 1/2 \cdot r\overset{1'}{\rho}), \quad \eta = (A^{-1}-1)/r^2 = 1/2 \cdot \overset{1}{\rho}, \quad \beta = \xi = 0, \qquad (2\cdot17)$$

and

$$\overset{1}{\rho} = 2\psi/r^2, \quad \overset{3}{\rho} = -4\psi/r^2, \qquad (2\cdot18)$$

respectively. An example of S_q is given by the space-time of Einstein universe whose line element in the r.c.c.s. is given by

$$ds^2 = -(1-r^2/R^2)^{-1}dr^2 - r^2(d\theta^2 + \sin^2\theta d\phi^2) + dt^2, \quad (\overset{1}{\rho} = -2/R^2), \qquad (2\cdot19)$$

where R^2 is a constant. We shall denote this S_q by $[$Ein. $(\overset{1}{\rho} = -2/R^2)]$, or simply by $[$Ein.$]$.

(iii) S_r i.e. an S which belongs to both S_p and S_q. For an S_r we have

$$\overset{1}{\rho} = -\overset{2}{\rho} = -\overset{3}{\rho} = 2\overset{4}{\rho} = 4m/\lambda^2, \qquad (2\cdot20)$$

where m is a constant, and its line element in the reduced canonical form is given by

as an S_p : $ds^2 = -(1+m)^{-1}dr^2 - r^2(d\theta^2 + \sin^2\theta d\phi^2) + (1+m)dt^2, \qquad (2\cdot21)$

as an S_q : $ds^2 = -(1+m)^{-1}dr^2 - r^2(d\theta^2 + \sin^2\theta d\phi^2) + dt^2. \qquad (2\cdot22)$

We shall denote this S_r by $S_r(m)$ when it is needed to indicate m. When $m \neq 0$, the

$S(m)$ is not Minkowski space-time [Min.], and it has two reduced canonical forms of the line element. Corresponding to (2·11) and (2·17), we have

$$\eta = 1/2 \cdot \overset{4}{\rho} = m/r^2, \quad \alpha = \beta = \xi = 0. \tag{2·23}$$

We shall call an S_p or an S_q which is not an S_r *proper*.

§ 3. Multiplication of a static S by a constant

Mathematically speaking, to introduce the concept of the superposition of the static S's we have only to make clear the following two operations: (i) Multiplication of a static S by a real constant p, and (ii) the summation of two static S's. In this section we shall define the first operation invariantly and shall obtain some results. Hereafter we shall denote various static S's by $(S)_1$, $(S)_2$, $\cdots$ and corresponding $\overset{a}{\rho}$'s $K_{ij}^{\cdot lm}$'s, X's, $\cdots$ by $(\overset{a}{\rho})_1$, $(\overset{a}{\rho})_2$, $\cdots$; $(K_{ij}^{\cdot lm})_1$, $(K_{ij}^{\cdot lm})_2$, $\cdots$; $(X)_1$, $(X)_2$, $\cdots$; $\cdots$ respectively.

Definition 1. *Let* $(S)_1$ *be given by* $(\overset{4}{\rho})_1$ *and* $(\overset{3}{\rho})_1$ *i.e.* $(S)^1 = f\{(\overset{4}{\rho})_1, (\overset{3}{\rho})_1\}$, *and p be a real constant, then the static S determined by*

$$(\overset{4}{\rho})_2 = p(\overset{4}{\rho})_1 \quad and \quad (\overset{3}{\rho})_2 = p(\overset{3}{\rho})_1 \tag{3·1}$$

is called the S obtained by multiplying the $(S)_1$ *by p.*

Accordingly if we denote the resulting S by $(S)_2$, we have

$$(S)_2 = p(S)_1 = (S)_1 p = f\{p(\overset{4}{\rho})_1, p(\overset{3}{\rho})_1\}. \tag{3·2}$$

When a quantity Q peculiar to an $(S)_1$ exists and the corresponding quantity of $p(S)_1$, where p is any constant, is given by pQ, then we shall say that the $(S)_1$ has the *linear property of the first kind* with respect to the Q. According to this terminology the above definition 1 is simply that we have defined the multiplication of $(S)_1$ by p using the *l. p. 1* (abbreviation of the linear property of the first kind) with respect to $\overset{4}{\rho}$ and $\overset{3}{\rho}$.

In the c.c.s. (3.1) is equivalent to

$$(\eta)_2 = p(\eta)_1, \qquad (\beta)_2 = p(\beta)_1. \tag{3·3}$$

Hence it is evident that when the $(S)_1$ is an S_p or an S_q or an S_r the resulting $(S)_2$ is again S_p or S_q or S_r respectively. Furthermore when the $(S)_1$ is the space-time of de Sitter or of Einstein or of Schwarzschild's exterior solution, the $(S)_2$ is again of the same type respectively. Moreover in all cases $(K_{ij}^{\cdot lm})_2$ is equal to $(K_{ij}^{\cdot lm})_1$ multiplied by p. When $p = 0$ the $(S)_2$ becomes [Min.] independently of the given $(S)_1$. [Min.] plays the rôle of 0 in the ordinary calculus by virtue of the relation [Min.]$= f(0, 0)$.

Theorem [3.1] *When* $(S)_2 = p(S)_1$, *it holds that*

$$(\overset{1}{\rho})_2 = p(\overset{1}{\rho})_1 + \varDelta, \quad (\overset{b}{\rho})_2 = p(\overset{b}{\rho})_1, \quad (b = 2, 3, 4), \tag{3·4}$$

where the 'interaction term' $\varDelta$ *is given by*

$$\varDelta = p(p-1)\lambda^2(X)_1(Y)_1/(Z)_1(Z)_2. \tag{3·5}$$

The proof is evident if we use the formulae given in the last section.

Hence any static S has the *l. p.* 1 with respect to $\overset{b}{\rho}$, $(b=2, 3, 4)$, though it is not the case with respect to $\overset{1}{\rho}$ in general. If we take

$$\overset{1}{\rho}{}^* = \overset{1}{\rho}\, Z - 1/4 \cdot \lambda^2 XY - 1/2 \cdot \lambda^3 \overset{1}{\rho}(\overset{3}{\rho}' + \overset{4}{\rho}') \tag{3.6}$$

in place of $\overset{1}{\rho}$, however, any static S has the *l. p.* 1 with respect to this $\overset{1}{\rho}{}^*$ also. The proof is evident from the fact that this $\overset{1}{\rho}{}^*$ is equal to $\lambda(\overset{3}{\rho}' + \overset{4}{\rho}')$ in reality. When the S is an S_p or an S_q the $\overset{1}{\rho}{}^*$ coincides with $\overset{1}{\rho}$.

From the results above obtained and the formulae in § 2, we have

Theorem [3.2] *When* $(S)_2 = p(S)_1$, *a necessary and sufficient condition that*

$$(\overset{a}{\rho})_2 = p(\overset{a}{\rho})_1, \qquad (a=1,\cdots,4) \tag{3.7}$$

hold is given by the condition that the given $(S)_1$ *be an* S_p *or an* S_q. (3.7) *is equivalent to*

$$(K_{ij}^{\cdot\cdot lm})_2 = p(K_{ij}^{\cdot\cdot lm})_1. \tag{3.8}$$

Accordingly any S_p *or* S_q *has the l. p.* 1 *with respect to* $\overset{a}{\rho}$, $(a=1,\cdots 4,)$, *and* $K_{ij}^{\cdot\cdot lm}$.

A static $(S)_1$ which is neither S_p nor S_q has not the *l. p.* 1 with respect to $K_{ij}^{\cdot\cdot lm}$. But using the fact that it has the *l. p.* 1 with respect to all components of $K_{ij}^{\cdot\cdot lm}$ except ξ in any c.c.s., we can express the relation between $(K_{ij}^{\cdot\cdot lm})_1$ and $(K_{ij}^{\cdot\cdot lm})_2$, the curvature tensor of $p(S)_2$, in tensor equation as follows:

For this purpose we shall consider a tensor $Q_{ij}^{\cdot\cdot lm}$ which satisfies

$$Q_{jilm} = -Q_{jilm} = -Q_{ijml} = Q_{lmij}, \tag{3.9}$$

and whose surviving components in any c.c.s. of any S is given by $Q_{14}^{\cdot\cdot 14} = 1/4$. Since this tensor is form-invariant under the transformation $(\bar{r}=r, \bar{t}=at+b)$, where a and b are constants, which transforms a c.c.s. into another c.c.s., this $Q_{ij}^{\cdot\cdot lm}$ is common to all static S's which we are considering.* From the theorem [3.1], we have $(\xi)_2 = p(\xi)_1 + 1/4 \cdot \varDelta$. Hence we obtain

$$\left\{ \begin{array}{l} (K_{ij}^{\cdot\cdot lm})_2 = p(K_{ij}^{\cdot\cdot lm})_1 + \varDelta Q_{ij}^{\cdot\cdot lm}, \\ (K_j^{\cdot l})_2 = p(K_j^{\cdot l})_1 + \varDelta Q_j^{\cdot l}, \quad (K_j^{\cdot l} = K_{ij}^{\cdot\cdot li}, \quad Q_j^{\cdot l} = Q_{ij}^{\cdot\cdot li}), \\ (K)_2 = p(K)_1 - 1/2 \cdot \varDelta, \quad (K = K_i^{\cdot i}), \end{array} \right. \tag{3.10}$$

* $Q_{ij}^{\cdot\cdot lm}$ as well as the canonical form of the line element is form-invariant under the group of spatial rotations. When the static S is of special kind, some transformations besides rotations may also keep the line element form-invariant. But by using the additional condition that the relation $\lambda = r$ must be kept invariant in the c.c.s., we can prove that it is not necessary to deal with these transformations except rotations to show that the $Q_{ij}^{\cdot\cdot lm}$ is common to all static S's which we are considering.

For any characteristic vectors of an $S^{5)}$, it holds that

$$Q_{ij}^{\cdot\cdot lm} = a_i a^l \beta_j \beta^m \tag{1}$$

to within an m-transformation at most, where a_i and β_i are any characteristic vectors of the S, since the right hand of (1) is form invariant under ε- and ω-transformations and the transformation $(\bar{a}_i = ha_i, \bar{\beta}_i = k\beta_i)$, (accordingly $\bar{a}^i = a^i/h, \bar{\beta}^i = \beta^i/k$). From (1) we have

$$Q_j^{\cdot l} = 1/4 \cdot (a_j a^l - \beta_j \beta^l), \quad Q_i^{\cdot i} = -1/2. \tag{2}$$

and accordingly, by putting $T_j^{\ l}=K_j^{\ l}-1/2\cdot K\delta_j^l$, we have

$$(T_j^{\ l})_2=p(T_j^{\ l})_1+\varDelta(Q_j^{\ l}-1/4\cdot\delta_j^l). \tag{3.11}$$

In a c.c.s. the non-vanishing components of $Q_j^{\ l}$ are given by $Q_1^{\ l}=Q_4^{\ l}=-1/4$.

From the fact that $4\bar{\bar{\xi}}$ is given by

$$\bar{\bar{\varDelta}}=1/4\cdot\lambda^2XYZ^{-1}+Y+\lambda Y' \tag{3.12}$$

in the c.c.s., it is evident that if we put

$$L_{ij}^{\ \ lm}=K_{ij}^{\ \ lm}-\bar{\bar{\varDelta}}Q_{ij}^{\ \ lm}, \tag{3.13}$$

then any static S has the *l. p.* 1 with respect to this $L_{ij}^{\ \ lm}$.

Lastly we shall express the relation $(S)_2=p(S)_1$ using the canonical forms of the line elements. Let the line elements of $(S)_h$, $(h=1, 2)$, in this form be

$$ds^2=-\{1+a_h(r)\}^{-1}dr^2-r^2(d\theta^2+\sin^2\theta d\phi^2)+C_h(r)dt^2. \tag{3.14}$$

Then from (2.8) we have

$$a_2=pa_1, \qquad C_2=\exp\int\frac{C'_1}{C_1}\frac{p(1+a_1)}{1+pa_1}dr. \tag{3.15}$$

Specially when the $(S)_1$ is an S_p or an S_q, using the reduced canonical forms we have

for $\qquad\qquad C_1=1+a_1, \qquad\qquad C_2=1+pa_1, \qquad$ (3.16)

or for $\qquad\qquad C_1=1, \qquad\qquad C_2=1,$ (3.17)

respectively.

Remark 1. Since $(3\cdot8)$ is a tensor equation, it holds in any coordinate system. It must be noticed, however, that since $(g_{ij})_2\neq(g_{ij})_1$, we have

$$(K_{ij}^{\ \ m})_2\neq p(K_{ij}^{\ \ m})_1, \quad (K_{ijlm})_2\neq p(K_{ijlm})_1, \text{ etc.} \tag{3.18}$$

Similar relations hold for (3.10) and (3.11).

Remark 2. When the $(S)_1$ is a proper S_p or S_q, we can define $p(S)_1$ by the one condition $(\overset{i}{\rho})_2=p(\overset{i}{\rho})_1$, and the *l. p.* 1 with respect to $\overset{ii}{\rho}$ follows as its natural consequence. Further if we define a scalar ν intrinsic to this $(S)_1$ by

$$\nu=1/2\cdot\lambda^2\overset{4}{\rho} \quad\text{and}\quad \nu=0 \tag{3.19}$$

according as the $(S)_1$ is a proper S_p and a proper S_q respectively, this ν becomes $C-1$ in the r.c.c.s. and the S has the *l. p.* 1 with respect to this ν i.e. it holds that $(\nu)_2=p(\nu)_1$.

§ 4. Summation of static S's

In this section we shall define the sum of static S's having their c.c.s. in common in the sense stated in § 2, and shall obtain some of its consequences.

Definition 2. *When*

$$(S)_1 = f\{ (\overset{4}{\rho})_1, (\overset{3}{\rho})_1 \}, \quad (S)_2 = f\{ (\overset{4}{\rho})_2, (\overset{3}{\rho})_2 \}, \tag{4·1}$$

we shall call the $(S)_3$ *defined by*

$$(\overset{4}{\rho})_3 = (\overset{4}{\rho})_1 + (\overset{4}{\rho})_2, \quad (\overset{3}{\rho})_3 = (\overset{3}{\rho})_1 + (\overset{3}{\rho})_2, \tag{4·2}$$

the sum of $(S)_1$ *and* $(S)_2$. *Namely it holds that*

$$(S)_3 = (S)_1 + (S)_2 = (S)_2 + (S)_1 = f\{ (\overset{4}{\rho})_1 + (\overset{4}{\rho})_2, (\overset{3}{\rho})_1 + (\overset{3}{\rho})_2 \}. \tag{4·3}$$

As in the last section if the quantity Q corresponding to $(S)_1$ and $(S)_2$ are $(Q)_1$ and $(Q)_2$ respectuvely and Q corresponding to $(S)_1 + (S)_2$ is $(Q)_1 + (Q)_2$, then we shall say that the S's have the linear property of the second kind (which is abbreviated as *l. p.* 2 hereafter) with respect to the quantity Q. Hence the above definition 2 requires the *l. p.* 2 with respect to $\overset{4}{\rho}$ and $\overset{3}{\rho}$.

By the definitions 1 and 2 we can consider the space-time of the form $p_1(S)_1 + p_2(S)_2 + \cdots$ where p's are arbitrary real constants. Again [Min.] plays the rôle of zero. Since in the c.c̣.s. (4.2) is equivalent to

$$(\eta)_3 = (\eta)_1 + (\eta)_2, \quad (\beta)_3 = (\beta)_1 + (\beta)_2, \tag{4·4}$$

we have the following theorem corresponding to [3.1]:

Theorem [4.1] *When* $(S)_3 = (S)_1 + (S)_2$, *it holds that*

$$(\overset{1}{\rho})_3 = (\overset{1}{\rho})_1 + (\overset{1}{\rho})_2 + \bar{\varDelta}, \quad (\overset{h}{\rho})_3 = (\overset{h}{\rho})_1 + (\overset{h}{\rho})_2, \quad (b = 2, 3, 4), \tag{4·5}$$

where the 'interaction term' $\bar{\varDelta}$ *is given by*

$$4 (Z)_1 (Z)_2 (Z)_3 \bar{\varDelta}/\lambda^2 = \{ (X)_1 (Y)_2 + (X)_2 (Y)_1 \} (Z)_1 (Z)_2 - 1/2 \cdot \lambda^2 (X)_1 (Y)_1 (Z)_2 (\overset{1}{\rho})_2$$

$$- 1/2 \cdot \lambda^2 (X)_2 (Y)_2 (Z)_1 (\overset{1}{\rho})_1, \tag{4·6}$$

and

$$(Z)_3 = 1 + 1/2 \cdot \lambda^2 (\overset{1}{\rho})_3 = (Z)_1 (Z)_2 - 1/4 \cdot \lambda^2 (\overset{1}{\rho})_1 (\overset{1}{\rho})_2. \tag{4·7}$$

Accordingly static S's have the *l. p.* 2 with respect to $\overset{b}{\rho}$, $(b = 2, 3, 4)$, but it is not the case with respect to $\overset{1}{\rho}$ in general. However, if we use the $\overset{1}{\rho}*$ introduced in the last section in place of $\overset{1}{\rho}$, we can easily see that the S's have the *l. p.* 2 with respect to this $\overset{1}{\rho}*$.

When $(\overset{1}{\rho})_3 = (\overset{1}{\rho})_1 + (\overset{1}{\rho})_2$ holds, we shall also say that the $(S)_1$ and $(S)_2$ are relatively conjugate concerning the *l. p.* 2 with respect to $\overset{1}{\rho}$ or shortly, $(S)_1$ and $(S)_2$ are relatively conjugate.

Theorem [4·2] *A necessary and sufficient condition that* $(S)_1$ *and* $(S)_2$ *be relatively conjugate is given by*

$$\bar{\varDelta} = 0, \tag{4·8}$$

and this is equivalent to the relation

$$(K_{ij}{}^{lm})_3 = (K_{ij}{}^{lm})_1 + (K_{ij}{}^{lm})_2. \tag{4·9}$$

The proof is evident from the fact that (4.8) is equivalent to $(\overset{a}{\rho})_3 = (\overset{a}{\rho})_1 + (\overset{a}{\rho})_2$, where $a = 1, \cdots, 4$. In the same way from (4.6) we have

Theorem [4.3] *Any two S_p's are relatively conjugate. The same holds for any two S_q's too.*

It is evident that when a static S is given, there exist infinitely many static S's which are relatively conjugate with the given S. Concerning this circumstance we can prove various theorems some of which we shall give in the following.

Theorem [4.4] 1°. *A proper S_p and a proper S_q are not relatively conjugate.* 2°. *[Min.] is relatively conjugate with all static S's. When $(S)_1$ is an S_r other than [Min.], a necessary and sufficient condition that $(S)_2$ is relatively conjugate with the $(S)_1$ is that the $(S)_2$ is an S_p or an S_q.* 3.° *Let $(S)_1$ be a proper S_p. If both $(S)_2$ and $(S)_3$ are not S_p and both are relatively conjugate with the $(S)_1$, then a necessary and sufficient condition that $p(S)_1 + q(S)_2$, where p and q are any real constants and the $p(S)_1 + q(S)_2$ is assumed not to be an S_p, be again relatively conjugate with the $(S)_1$ is given by $p + q = 1$. The same holds also even if we use S_q in place of S_p.*

Proof. 1.° If $(S)_1$ and $(S)_2$ are S_p and S_q respectively, we have $\bar{J} = (X)_2 (Y)_1 (Z)_3^{-1} \cdot \lambda^2/4$. Hence the proof is evident. 2° and 3° can also be proved easily.

If $(S)_3 = (S)_1 + (S)_2$, we have the following relations corresponding to (3.10) (3.11):

$$\begin{cases} (K_{ij}{}^{lm})_3 = (K_{ij}{}^{lm})_1 + (K_{ij}{}^{lm})_2 + \bar{J} Q_{ij}{}^{lm}, \\ (K_j{}^l)_3 = (K_j{}^l)_1 + (K_j{}^l)_2 + \bar{J} Q_j{}^l, \\ (K)_3 = (K)_1 + (K)_2 - 1/2 \cdot \bar{J}, \end{cases} \qquad (4 \cdot 10)$$

$$(T_j{}^l)_3 = (T_j{}^l)_1 + (T_j{}^l)_2 + \bar{J}(Q_j{}^l - 1/4 \cdot \delta_j{}^l), \qquad (4 \cdot 11)$$

and we have the *l. p.* 2 with respect to $L_{ij}{}^{lm}$ defined by (3.13) i.e.

$$(L_{ij}{}^{lm})_3 = (L_{ij}{}^{lm})_1 + (L_{ij}{}^{lm})_2. \qquad (4 \cdot 12)$$

Here, it is needless to say that similar remark to the remark 1 of § 3 is also necessary concerning these equations, and further when both $(S)_1$ and $(S)_2$ are S_p or S_q a remark corresponding to the remark 2 of § 3 is also necessary with a proviso that we must use $(\nu)_3 = (\nu)_1 + (\nu)_2$ in place of $(\nu)_2 = p(\nu)_1$.

Next as in § 3, we shall express the relation $(S)_3 = (S)_1 + (S)_2$ using the canonical forms of their line elements. Let the line elements of $(S)_b$, $(b = 1, 2, 3)$, be

$$ds^2 = -\{1 + a_b(r)\}^{-1} dr^2 - r^2(d\theta^2 + \sin^2\theta d\phi^2) + C_b(r) dt^2, \quad (b = 1, 2, 3), \qquad (4 \cdot 13)$$

then from (2.8) we can easily obtain

$$a_3 = a_1 + a_2,$$

$$C_3 = \exp \int \left\{ (1 + a_1) \frac{C'_1}{C_1} + (1 + a_2) \frac{C'_2}{C_2} \right\} (1 + a_1 + a_2)^{-1} dr. \qquad (4 \cdot 14)$$

Specially when both $(S)_1$ and $(S)_2$ are S_p, using the reduced canonical forms, we have

$$C_3 = 1 + a_1 + a_2 \text{ for } C_1 = 1 + a_1, \; C_2 = 1 + a_2. \qquad (4 \cdot 15)$$

Similarly when both $(S)_1$ and $(S)_2$ are S_q, we have $C_b = 1$ for $b = 1, 2, 3$.

Hitherto we have introduced the two operations $p(S)_1$ and $(S)_1 + (S)_2$ and studied their properties. By this fact we may say that we have succeeded in introducing the concept of the superposition of the static spherically symmetric space-times. Now we shall denote by $\mathfrak{S}_p$ the set of all S_p's and by $\mathfrak{S}_q$ that of all S_q's. It has been made clear that the elements of $\mathfrak{S}_p$ (or $\mathfrak{S}_q$) have both $l.\,p.\,1$ and $l.\,p.\,2$ with respect to $\overset{a}{\rho}$, $(a = 1, \cdots, 4)$, (accordingly with respect to K_{ij}^{lm}), for these two operations. This is the main reason why we adopted the definitions 1 and 2.

Lastly we shall give a word about a type of S's frequently dealt with in general relativity. Here let us consider an S depending on some continuous parameters m_1, m_2, $\cdots$ and denote it by $S(m_1, m_2, \cdots)$. [de S.], [Ein.] and [Sch. ex.] are of this type. If such an S is an S_p or an S_q and its $\overset{1}{\rho}$ and $\overset{?}{\rho}$ are linear homogeneous functions of these parameters as is seen in these examples, we have evidently

$$pS(m_1, m_2, \cdots) + qS(\overline{m}_1, \overline{m}_2, \cdots) = S(pm_1 + q\overline{m}_1, pm_2 + q\overline{m}_2, \cdots), \quad (4\cdot16)$$

where p and q are arbitray constants.

§ 5   Some theorems concerning the superposition of the static S's

In this section we shall give some of the other theorems concerning the super-position of the static S's.

Theorem [5.1]   *If a set $\mathfrak{S}$ of static S's has the following two properties, then the $\mathfrak{S}$ is $\mathfrak{S}_p$ or $\mathfrak{S}_q$ or a sub-set of either of these two sets.*

1°.   *Any element $(S)_1$ of $\mathfrak{S}$ has the $l.\,p.\,1$ with respect to $\overset{1}{\rho}$, and $p(S)_1$, where p is any real constant, belongs to the $\mathfrak{S}$ again.*

2°.   *Any two elements $(S)_1$ and $(S)_2$ of $\mathfrak{S}$ have the $l.\,p.\,2$ with respect to $\overset{1}{\rho}$ and the sum $(S)_1 + (S)_2$ belongs to the $\mathfrak{S}$ again.*

Proof.   Let any element of the $\mathfrak{S}$ be $(S)_1$. From [3.2] $(S)_1$ must be either an S_p or an S_q. If $(S)_1$ is an S_p (or an S_q) and $(S)_2$ is any element of $\mathfrak{S}$ other than $(S)_1$, then it must be that $J = 0$ for $(S)_1$ and $p(S)_2$ and from this relation we can easily show that the $(S)_2$ must be an S_p (or an S_q) again. Hence the theorem is obvious.

Theorem [5.2]   *Let $(S)_1$ and $(S)_2$ be an S_p and an S_q respectively. A necessary and sufficient condition that $(S)_3 = (S)_1 + (S)_2$ be an S_p (or an S_q) is given by the condition that $(S)_2$ (or $(S)_1$) be an S_p. Accordingly when both $(S)_1$ and $(S)_2$ are not S_p, the sum $(S)_3$ is neither an S_p nor an S_q.*

Proof.   If $(S)_3$ is an S_p, from

$$(\overset{..}{\rho})_3 = (\overset{..}{\rho})_1 + (\overset{..}{\rho})_2 = (\overset{3}{\rho})_1 - 2(\overset{4}{\rho})_2 = (\overset{2}{\rho})_3 = \lambda(\overset{4}{\rho'})_3 = (\overset{2}{\rho})_1 + (\overset{2}{\rho})_2, \quad (5\cdot1)$$

we have $(\overset{3}{\rho})_2 = -2(\overset{4}{\rho})_2 = (\overset{3}{\rho})_2$. Hence $(S)_2$ is an S_p. The converse is evident. Also when $(S)_3$ is an S_q, we can prove similarly.

$\overset{4}{\rho}$ and $\overset{3}{\rho}$ of the sum of an $S = f\{\overset{4}{\rho}, \overset{3}{\rho}\}$ and $S_r(m)$ are given by

$$\overset{4}{\rho} + 2m/\lambda^2, \qquad \overset{3}{\rho} - 4m/\lambda^2, \quad (5\cdot2)$$

and the corresponding $\bar{\Delta}$ is given by

$$\bar{\Delta} = -1/4 \cdot \lambda^2 \, mXY/Z(Z+m). \tag{5.3}$$

If the S is an S_p or an S_q, $\Delta=0$ evidently. Using these relations we can prove the following decomposition theorem:

Theorem [5.3] *Any static S can be decomposed as a sum of an S_p and an S_q. In this decomposition the S_p is determined uniquely to within an arbitrary additive $S_r(m)$ and consequently the S_q is also determined uniquely to within* $-S_r(m)$.

Proof. In the equation $S = (S)_1 + (S)_2$, if $(S)_1$ and $(S)_2$ are S_p and S_q respectively, from (4.2) we have

$$\overset{3}{\rho} = \lambda(\overset{4}{\rho'})_1 - 2(\overset{4}{\rho})_2, \quad \overset{4}{\rho} = (\overset{4}{\rho})_1 + (\overset{4}{\rho})_2, \tag{5.4}$$

from which we get

$$(\overset{4}{\rho})_1 = P, \quad (\overset{4}{\rho})_2 = \overset{4}{\rho} - P, \tag{5.5}$$

where

$$P = \lambda^{-2} \{ \int (\overset{3}{\rho} + 2\overset{4}{\rho}) \lambda d\lambda + 2m \}, \tag{5.6}$$

and m is any constant. The corresponding $\overset{3}{\rho}$'s and $\overset{2}{\rho}$'s are given by

$$(\overset{3}{\rho})_1 = (\overset{2}{\rho})_1 = \lambda(\overset{4}{\rho'})_1 = \overset{3}{\rho} + 2\overset{4}{\rho} - 2P, \tag{5.7}$$

$$(\overset{3}{\rho})_2 = -2(\overset{4}{\rho})_2 = -2\overset{4}{\rho} + 2P.$$

From these results the theorem is obvious.

In this theorem, when S is an S_p, from

$$\lambda(\overset{3}{\rho} + 2\overset{4}{\rho}) = \lambda(\overset{2}{\rho} + 2\overset{4}{\rho}) = (\lambda^2 \, \overset{4}{\rho})', \tag{5.8}$$

we have $(\overset{1}{\rho})_2 = -1/2 \cdot (\overset{3}{\rho})_2 = -2m/\lambda^2$. Hence $(S)_2$ is an S_r. Similarly when S is an S_q, $(S)_1$ becomes an S_r. These results are consistent with [5.2].

From [5.3] it follows that if two decompositions of S be

$$S = (S)_1 + (S)_2 = (\overline{S})_1 + (\overline{S})_2, \tag{5.9}$$

then there exists an $S_r(m)$ satisfying

$$(\overline{S})_1 = (S)_1 + S_r(m), \quad (\overline{S})_2 = (S)_2 - S_r(m). \tag{5.10}$$

We shall call the $(S)_1$ and $(S)_2$ in [5.3] the S_p- and S_q-components of the given S.

If we express the decomposition theorem by using the c.c.s. it becomes as follows:

Let ds^2 of the S in this coordinate system be

$$ds^2 = -\{1 + \psi(r)\}^{-1} dr^2 - r^2(d\theta^2 + \sin^2\theta d\phi^2) + C(r)dt^2, \tag{5.11}$$

and the reduced canonical forms of ds^2's, of its S_p- and S_q-components be

$$ds^2 = -\{1 + a(r)\}^{-1}dr^2 - r^2(d\theta^2 + \sin^2\theta d\phi^2) + \{1 + a(r)\}dt^2, \tag{5.12}$$

and

$$ds^2 = -\{1 + b(r)\}^{-1} dr^2 - r^2(d\theta^2 + \sin^2\theta d\phi^2) + dt^2, \tag{5.13}$$

respectively. Then a and b are given by

$$a=\int\{C'(1+\psi)/C\}\,dr+m,\quad b=\psi-a,\qquad\qquad (5\cdot14)$$

where m is an arbitrary constant and gives the additive $S_r(m)$ in [5.3].

Concerning this result it is also to be noticed that if $K_{ij}{}^{,lm}$'s, $T_i{}^{,p}$'s, etc. are known for $(S)_1$ and $(S)_2$, then those for $(S)_1$ and $\overline{(S)}_2$ are easily obtained by the remark given after [5.2].

Theorem [5.4] *A necessary and sufficient condition that the S_p-component of a static S be* [de S. $(\overset{\shortmid}{\rho}=-2k^2)$] *is given by* $Y=\overset{..}{\rho}+2\overset{\shortmid}{\rho}=-4k^2$, *and the condition that the S_q-component be* [Ein. $(\overset{4}{\rho}=-2/R^2)$] *is given by* $X=\overset{..}{\rho}-\lambda\overset{\shortmid}{\rho}=4/R^2$.

The proof is easy, so we shall omit it here. It is needless to say that specially when $k^2=0$ or $1/R^2=0$, the S must be S_q or S_p respectively.

Similar theorems can also be obtained if we use some special S_p or S_q in place of [de S.] or [Ein.] respectively.

Lastly we shall add a theorem concerning a static S which is conformal to an S_{20}.*

Theorem [5.5] *Let $(S)_1$ be conformal to an S_{20}. If $(S)_1$ has the l. p. 1 with respect to ρ, then $p(S)_1$ is also conformal to an S_{20}. Similarly when both $(S)_1$ and $(S)_2$ are conformal to some S_{20}'s and are relatively conjugate, then $(S)_1+(S)_2$ is again conformal to an S_{20}. The same statement holds if we use 'conformally flat' in place of 'conformal to an S_{20}'.*

Proof. The theorem is evident because a necessary and sufficient condition that an S be conformal to an S_{20} is $\overset{\shortmid}{\rho}\lambda^2=$const., and the condition that an S be conformally flat is $\overset{\shortmid}{\rho}=0$.[5]

From both this theorem and the fact that an S_r is conformal to an S_{20} we know that the sum of an S_p (or an S_q) conformal to an S_{20} and an S_r is again conformal to an S_{20}. But the sum of a conformally flat S_p(or S_q) and an $S_r(m)$, $(m\neq0)$, is not conformally flat.

As is seen in the theorems obtained in this section, when some intrinsic property of static S's are expressed in terms of ρ's and these expressions are linear and homogeneous when expressed as functions of $\overset{\shortmid}{\rho}$ and $\overset{..}{\rho}$, the property is preserved by the operations $p(S)_1$ and $(S)_1+(S)_2$. But when such a circumstance is not the case we must put some additional assumptions e.g., linearity with respect to $\overset{\shortmid}{\rho}$, to preserve the property by these operations. To illustrate this we shall give an example: The condition for an S to be of class one is given by $\overset{\shortmid}{\rho}+0$ and $2\overset{\shortmid}{\rho}\overset{..}{\rho}=\overset{\shortmid}{\rho}\overset{..}{\rho}$.[7] Under the additional assumption of the l. p. 1 with respect to $\overset{\shortmid}{\rho}$, this property is preserved by the operation $p(S)_1$. As to the summation of two S's, however, we can not obtain such a simple result.

§ 6. Examples of the superposition and the decomposition of static S's

(i) As is easily seen from the discussion at the end of § 3, the S's obtained by multiplying

* The concept of an S conformal to an S_{20} was introduced by the writer by generalizing that of conformally flat.[6]

[de S.], [Ein.] and [Sch. ex.] by p are of the same type as the original ones, and we have

$$p[\text{de S.} (\overset{4}{\rho} = -2k^2)] = [\text{de S.} (\overset{4}{\rho} = -2k^2 p)],$$

$$p[\text{Ein.} (\overset{4}{\rho} = -2/R^2)] = [\text{Ein.} (\overset{4}{\rho} = -2p/R^2)],$$

$$p[\text{Sch. ex.} (\overset{4}{\rho} = -4m\lambda^{-3})] = [\text{Sch. ex.} (\overset{4}{\rho} = -4pm\lambda^{-3})].$$

(ii) For the space-time S_p obtained as the sum of [de S. $(\overset{4}{\rho} = -2k^2)$] and [Sch. ex. $(\overset{4}{\rho} = -4m\lambda^{-3})$], we have

$$\overset{1}{\rho} = -2\overset{2}{\rho} = -2\overset{3}{\rho} = -24m\lambda^{-3}, \quad \overset{4}{\rho} = -2k^2 - 4m\lambda^{-3}. \tag{6·1}$$

Accordingly the reduced canonical form of its line element is given by

$$ds^2 = -(1 - k^2 r^2 - 2m/r)^{-1} dr^2 - r^2 (d\theta^2 + \sin^2 \theta d\phi^2) + (1 - k^2 r^2 - 2m/r) dt^2. \tag{6·2}$$

Non-vanishing components of $K_{ij}{}^{lm}$ for this ds^2 are given by

$$\alpha = \beta = -k^2 + m/r^3, \quad \xi = \eta = -k^2 - 2m/r^3, \tag{6·3}$$

and (4.9) holds. In the usual theory of general relativity this S_p defined by (6.2) is obtained as the static spherically symmetric solution of the field equation with the cosmological term $K_{ij} = 3k^2 g_{ij}$, but it has never been obtained, it seems to the writer, as the result of the superposition of two space-times.

This (ii) is an example of the sum of two S_p's.

(iii) Next, as an example of the summation of two S_q's, we shall consider the sum of [Ein. $(\overset{4}{\rho} = -2/R^2)$] and the S_q defined by

$$ds^2 = -(1 + mr)^{-1} dr^2 - r^2 (d\theta^2 + \sin^2 \theta d\phi^2) + dt^2, \tag{6·4}$$

where m is a constant. Since

$$-2\overset{1}{\rho} = 2\overset{2}{\rho} = \overset{3}{\rho} = -2\overset{4}{\rho} = -4m/\lambda \tag{6·5}$$

for (6·4), we have

$$\overset{1}{\rho} = -\overset{2}{\rho} = 2m/\lambda, \quad \overset{3}{\rho} = -2\overset{4}{\rho} = 4/R^2 - 4m/\lambda \tag{6·6}$$

for the resulting S_q, and the reduced canonical form of its line element is given by

$$ds^2 = -(1 + mr - r^2/R^2)^{-1} dr^2 - r^2 (d\theta^2 + \sin^2 \theta d\phi^2) + dt^2. \tag{6·7}$$

As a matter of course (4.9) also holds in this case.

(iv) For the sum [de S. $(\overset{4}{\rho} = -2k^2)$] + [Ein. $(\overset{4}{\rho} = -2/R^2)$], we have

$$\overset{1}{\rho} = -(4k^2\lambda^2/R^2)\{1 - (k^2 + 1/R^2)\lambda^2\}^{-1}, \quad \overset{2}{\rho} = 0,$$

$$\overset{3}{\rho} = 4/R^2, \quad \overset{4}{\rho} = -2(k^2 + 1/R^2),$$

and this S is neither S_p nor S_q. The line element of this space-time takes the following form in the c.c.s.:

$$ds^2 = -(1 - k^2 r^2 - r^2/R^2)^{-1} dr^2 - r^2 (d\theta^2 + \sin^2 \theta d\phi^2) + C dt^2, \tag{6·9}$$

where
$$C = \text{const.} \times \{1 - (k^2 + 1/R^2) r^2\}^{k^2/(k^2+1/R^2)}.$$

(v)   For the sum [Ein. $(\overset{4}{\rho} = -2/R^2)$] + [Sch. ex. $(\overset{4}{\rho} = -4m/\lambda^3)$], we have

$$
\left\{
\begin{array}{l}
\overset{1}{\rho} = \dfrac{4m}{R^2\lambda(1 - \lambda^2/R^2 - 2m/\lambda)} - \dfrac{24m}{\lambda^3}, \quad \overset{2}{\rho} = \dfrac{12m}{\lambda^3}, \quad \overset{3}{\rho} = \dfrac{4}{R^2} + \dfrac{12m}{\lambda^3}, \\[3mm]
\overset{4}{\rho} = -\dfrac{2}{R^2} - \dfrac{4m}{\lambda^3},
\end{array}
\right.
\tag{6.10}
$$

and again this S is neither S_p nor S_q. ds^2 in the c.c.s. is given by

$$ds^2 = -(1 - r^2/R^2 - 2m/r)^{-1} dr^2 - r^2(d\theta^2 + \sin^2\theta d\phi^2) + Cdt^2, \tag{6.11}$$

where
$$C = \text{const.} \times \exp \int 2m \{r^2(1 - r^2/R^2 - 2m/r)\}^{-1} dr.$$

(vi)   For the space-time of Schwarzschild's interior solution whose line element in a c.c.s. is given by

$$ds^2 = -(1 - r^2/R^2)^{-1} dr^2 - r^2(d\theta^2 + \sin^2\theta d\phi^2) + (a - b\sqrt{1 - r^2/R^2})^2 dt^2, \tag{6.12}$$

where $1/R^2$, a and b are constants, we have

$$\overset{1}{\rho} = \overset{2}{\rho} = 0, \quad \overset{3}{\rho} = (4a/R^2)(a - b\sqrt{1 - r^2/R^2})^{-1}, \quad \overset{4}{\rho} = -2/R^2. \tag{6.13}$$

The S_p- and S_q-components of this S are given by

$$S_p\text{-component}: \qquad \overset{4}{\rho} = 2P\lambda^{-2}, \quad \overset{3}{\rho} = 2\lambda^{-1} P' - 4\lambda^{-2} P, \tag{6.14}$$

$$S_q\text{-component}: \qquad \overset{4}{\rho} = -2R^{-2} - 2\lambda^{-2}P, \quad \overset{3}{\rho} = 4R^{-2} + 4\lambda^{-2} P,$$

where $P(\lambda) = (1 - \lambda^2/R^2) + (2a/b)\sqrt{1 - \lambda^2/R^2} + (2a^2/b^2)\log(a - b\sqrt{1 - \lambda^2/R^2})$, (6.15)

to within an arbitrary additive $S_r(m)$. The line elements of both components in the respective r.c.c.s. are given by (5.12) and (5.13) in which $a = P(r)$ and $b = -r^2/R^2 - P(r)$ respectively.

§ 7.   Multiplication, division, etc. of static S's

In the previous sections we introduced two kinds of operation $p(S)_1$ and $(S)_1 + (S)_2$ concerning static S's, studied its consequences, and further gave some examples. Now in this section we shall show that we can also introduce the operations $(S)_1 (S)_2$, $(S)_1/(S)_2$, $d/d\lambda \cdot (S)_1$, etc. consistently with the above two. This research is made mainly from a mathematical interest and may, at the present stage, have little significance physically.

As is seen from (2.2) there exists a one to one correspondence between static S's and pairs of real functions $\overset{1}{\rho}(\lambda)$ and $\overset{3}{\rho}(\lambda)$. From this we can deduce that if we denote by $x(\lambda)$ and $y(\lambda)$ any two linear combinations of $\overset{4}{\rho}(\lambda)$ and $\overset{3}{\rho}(\lambda)$ with real constant coefficients and linearly independent relatively to each other, then we get a new one to one correspondence between the S's and $x(\lambda) + iy(\lambda)$. Hence we can define the addition, substraction, multiplication, division, etc. of the S's by using the corresponding operations concerning such complex functions. If we explain this more in detail by using mathematical

expressions it becomes as follows :

Let $\{(x)_1, (y)_1\}$ and $\{(x)_2, (y)_2\}$ be two pairs of x and y corresponding to $(S)_1$ and $(S)_2$ and p and q be any constants, then $p(S_1) + q(S)_2$, $(S)_1(S)_2$, $(S)_1 \div (S)_2$, $d/d\lambda \cdot (S)_1$, $\cdots$ are, by definition, the S's corresponding to $\{p(x)_1 + q(x)_2, p(y)_1 + q(y)_2\}$, $\{(x)_1(x)_2 - (y)_1(y)_2, (x)_1(y)_2 + (x)_2(y)_1\}$, $[\{(x)_1(x)_2 + (y)_1(y)_2\} \{(x)_2^2 + (y)_2^2\}^{-1}$, $\{(x)_2(y)_1 - (x)_1(y)_2\} \{,, \quad ,,\}^{-1}]$, $\{(x)_1', (y)_1'\}, \cdots$, respectively. It is evident that the operations thus defined are consistent with those introduced in § 3 and § 4, and that they satisfy the fundamental laws of arithmetics and differential and integral calculus. This will enable us to discuss about the functions of the S's. In these operations the space-time which corresponds to the origin of such a functional space, i.e. the point $x = y = 0$, is given by the Minkowski space-time [Min.]. Hence division by [Min.] is impossible.

Next we shall consider the problem : What kinds of linear combination of $\overset{4}{\rho}$ and $\overset{3}{\rho}$ are to be taken as $x(\lambda)$ and $y(\lambda)$? As is evident from the property of the complex plane, we can determine so as any two linearly independent static S's whose $\overset{4}{\rho}$ and $\overset{3}{\rho}$ are both constant may correspond to the purely real and purely imaginary units. In the following, as an example, we shall give a definition of x and y by which the space-time [de S. $(\overset{4}{\rho} = -2k^2)$] corresponds to the real number k^2 and the space-time [(Ein. $(\overset{4}{\rho} = -2/R^2)$] to the imaginary number i/R^2.

For this purpose we have only to put

$$x = -1/4 \cdot (\overset{3}{\rho} + 2\overset{4}{\rho}), \quad y = 1/4 \cdot \overset{3}{\rho}. \tag{7.1}$$

Evidently this correspondence has the property above stated.* Then any S_q corresponds to a purely imaginary functions. Examples of the correspondence are given by

$$[\text{de S.}(\overset{4}{\rho} = -2(k^2)] \longleftrightarrow (k^2, 0), \quad [\text{Ein. } (\overset{4}{\rho} = -2/R^2)] \longleftrightarrow (0, 1/R^2),$$

$$[\text{Sch. ex. } (\overset{4}{\rho} = -4m\lambda^{-3})] \longleftrightarrow (-m\lambda^{-3}, 3m\lambda^{-3}), \quad S_r(m) \longleftrightarrow (0, -m\lambda^{-2}), \tag{7.2}$$

$$(S)' \longleftrightarrow (x', y'), \quad \textstyle\int S d\lambda \longleftrightarrow (\int x d\lambda, \int y d\lambda).$$

Further we shall give some examples of the multiplication, differentiation, etc., applied to some special S's :

$$[\text{Ein.}(\overset{4}{\rho} = -2/R^2)] \times [\text{de S.}(\overset{4}{\rho} = -2k^2)] = [\text{Ein.}(\overset{4}{\rho} = -2k^2/R^2)],$$

$$[\text{Sch. ex.}(\overset{4}{\rho} = -4m\lambda^{-3})] \times [\text{de S.}(\overset{4}{\rho} = -2k^2)] = [\text{Sch. ex.}(\overset{4}{\rho} = -4k^2 m\lambda^{-3})], \tag{7.3}$$

$$[\text{Sch. ex.}(\overset{4}{\rho} = -4m\lambda^{-3})] \times [\text{Ein.}(\overset{4}{\rho} = -2/R^2)]$$

$$= \text{static } S \text{ defined by } \{x = -3m/R^2\lambda^3, \ y = -m/R^2\lambda^3\}$$

$$= \quad ,, \qquad ,, \qquad \{\overset{4}{\rho} = 8m/\lambda^3 R^2, \ \overset{3}{\rho} = -4m/\lambda^3 R^2\}, \tag{7.4}$$

---

* Since the relation (7.1) is expressed by $x = -\beta$ and $y = \beta - \eta$ in the c.c.s., the equation $(S)_3 = (S)_1 \cdot (S)_2$ is equivalent to

$$-(\beta)_3 = (\beta)_1(\eta)_2 + (\beta)_2(\eta)_1 - (\eta)_1(\eta)_2,$$

$$-(\eta)_3 = 2\{(\beta)_1(\eta)_2 + (\beta)_2(\eta)_1\} - 2(\beta)_1(\beta)_2 - (\eta)_1(\eta)_2.$$

whose line element in a c.c.s. is given by

$$ds^2 = -(1+4m/R^2 r)^{-1} dr^2 - r^2(d\theta^2 + \sin^2\theta d\phi^2) + (1+4m/R^2 r)^{-3/2} dt^2, \qquad (7\cdot5)$$

$$\frac{d}{d\lambda}[\text{de S.}] = \frac{d}{d\lambda}[\text{Ein.}] = 0, \qquad (7\cdot6)$$

$$\int[\text{de S.}(\overset{4}{\rho}=-2k^2)]d\lambda = \text{static } S \text{ defined by } \{x=k^2\lambda+a, y=b\}$$
$$= \quad \text{,,} \qquad \text{,,} \qquad \{\overset{4}{\rho}=-2(k^2\lambda+a+b), \overset{3}{\rho}=4b\}, \qquad (7\cdot7)$$

where a and b are constants.

When $(S)_3 = (S)_1(S)_2$ holds, the relation among the curvature tensors of the three S's is not simple in general as is seen from the footnote concerning $(7\cdot1)$. However, specially when the $(S)_1$ is [de S.$(\overset{4}{\rho}=-2k^2)$] and the $(S)_2$ is an S_p or an S_q, it holds that

$$(K_{ij}^{\cdot lm})_3 = k^2(K_{ij}^{\cdot lm})_2. \qquad (7\cdot8)$$

When we take functions which are not linear combinations of $\overset{4}{\rho}$ and $\overset{3}{\rho}$ as $x(\lambda)$ and $y(\lambda)$, we have somewhat different results. As an example we shall put

$$x(\lambda) = -1/4 \cdot Y = -1/4 \cdot (\overset{3}{\rho}+2\overset{4}{\rho}), \quad y(\lambda) = 1/4 \cdot X = 1/4 \cdot (\overset{3}{\rho}-\lambda\overset{4}{\rho}'). \qquad (7\cdot9)$$

Then S_p's and S_q's correspond to $(x, 0)$ and $(0, y)$ respectively. And the present correspondence differs greatly from the former one in that the correspondence between (x, y) and S's is not one to one. Speaking further, though one pair of x and y corresponds to any S, innumerable static S's correspond to one pair of x and y with an arbitrary additive $S_r(m)$, because by solving (7.9) conversely, we have

$$\overset{4}{\rho} = -\frac{4}{\lambda^2}\int(x+y)\lambda d\lambda + \frac{2m}{\lambda^2}, \quad \overset{3}{\rho} = -4x + \frac{8}{\lambda^2}\int \quad \text{,,} \quad d\lambda - \frac{4m}{\lambda^2}, \qquad (7\cdot10)$$

where m is an arbitrary constant and gives the $S_r(m)$.

Accordingly, in this case, operations concerning S's correspond to those concerning $(x+iy)$'s under the additional condition 'mod $S_r(m)$', e.g., $(0, 0)$ corresponds to $\{[\text{Min.}] +$ arbitrary $S_r(m)\}$, real number k^2 to $\{[\text{de S.}(\overset{4}{\rho}=-2k^2)]+$ arbitrary $S_r(m)\}$.

In this section we have introduced various operations concerning static S's. It is evident, however, that further we can introduce the concept of space-time valued functions of space-times $F\{(S)_1, (S)_2, \cdots\}$, their series expansion, differentiation by space-times, etc.

§8. Conclusion

We have emphasized the possibility of solving the problem of many bodies in general relativity by introducing the concept of the superposition of space-times and have constructed an invariant theory of the superposition concerning static spherically symmetric space-times having the centers of symmetry in common. This theory, in the present stage, is insufficient to apply to the problem of many bodies and it is particularly desirable to establish the theory for the space-times which have different centers of symmetry. And

if we only succeed in establishing this we may expect the solution of the problem of many bodies.

Here we must notice again that the solution of the problem by this method will not necessarily lead to the complete solution in the sense hitherto considered. Similar circumstances have been seen in the history of the development of the quantum mechanics ; the stage of solving the problem of many bodies by using the superposition of space-times may correspond to that of solving classical equations of motion under some quantum conditions in the early stage of the quantum theory.

The writer is grateful to Professor Y. Mimura for his helpful discussions.

§ 9. Remark—The problem of cosmical effect to a space-time

Our theory has been constructed with the object of providing a method of solving the problem of many bodies. In this section we intend to show that this theory is also applicable to another problem in general relativity.

From the examples in § 6, we can see that we have come to deal with the *cosmical effect* i.e. the effect due to the whole universe, to a spherically symmetric space-time. For this purpose we have only to consider the superposition of the given space-time and the model of universe we wish to consider. In the following we shall deal somewhat in detail with this byproduct of our theory.

First we shall start with the question : What space-time is obtained if we consider the cosmical effect upon the space-time of Schwarzschild's exterior solution ? In the usual theory of relativity the answer is given by the space-time defined by (6.2) by solving the gravitational equation with the cosmical term $K_{ij} = 3k^2 g_{ij}$ in place of (1.1). According to the present theory, however, this space-time is nothing but the one obtained by superposing the Schwarzschild's space-time and the de Sitter universe. Furthermore if we use the present theory we can consider another kind of the cosmical effect. For example, we can superpose the Schwarzschild's space-time and the Einstein universe as was seen in (v) of § 6.

Now we shall give another example : In the usual theory of relativity, in order to determine the space-time which expresses the inner structure of some body we express mathematically the characteristic of the given structure as the conditions to be satisfied by the energy-momentum tensor T_{ij} and solve the field equation

$$K_{ij} - 1/2 \cdot K g_{ij} = 8\pi T_{ij}. \qquad (9 \cdot 1)$$

Further if we intend to consider the cosmical effect upon this space-time we solve the field equation with the cosmical term

$$K_{ij} - 1/2 \cdot K g_{ij} + \Lambda g_{ij} = 8\pi T_{ij}, \qquad (9 \cdot 2)$$

in place of (9.1). However, by virtue of the present theory, we have a new method of solving this problem by superposing the given space-time and any universe we wish to consider. It is to be noticed, however, that the space-time with the cosmical effect thus obtained does not necessarily satisfy the field equation (9.2). This situation is similar to

that which occurred concerning the problem of many bodies.

We have seen that the theory of the superposition of space-times is applicable to the problem of the cosmical effect. However, in the present stage of our theory, we cannot consider any non-static cosmical effect in spite of the fact that the models of the universe in relativistic cosmology are non-static in general.

The writer hopes that as the theory of the superposition of space-times advances it will be of use in solving not only the problems treated in this paper but also many other problems in the general theory of relativity.

References

1) L. Silberstein, Phys. Rev. **49**(1936), 208.
2) T. Hosokawa, Journ. Sci. Hiroshima Univ. A, **8**(1938), 29.
3) A. S. Eddington, *The Mathematical Theory of Relativity*, Cambridge, (1924), 101.
4) H. Takeno, Prog. Theor. Phys. **10**(1953), 509. This paper is cited as [**4**] in the text.
5) H. Takeno, Journ. Math. Soc. Japan **3**(1951), 317.
6) H. Takeno, Journ. Sci. Hiroshima Univ. A, **16**(1952), 299.
7) H. Takeno, Journ. Sci. Hiroshima Univ. A, **16**(1952), 291.

A Tentative Theory of Λ-Particles
— *Theory of 4-dimensional τ-space* —

Tadao NAKANO and Ryôyû UTIYAMA

Osaka City University and Osaka University

In order to describe new kinds of particles a four dimensional Euclidian τ-space S_4 is introduced. Every elementary particle is supposed to be described by a particular irreducible spinor in S_3 which is imbedded in S_4 as a subspace. Any one of wave functions is assumed to correspond respectively to a particular one of irreducible spinors in S_4, which will be decomposed into a sum of several parts, each one of which corresponds to a particular kind of elementary particles. In the present paper the nucleon and the Λ-particle are described by a single spinor corresponding to the representation $R_{1/2,1/2}(=\mathfrak{D}_{1/2,0}+\mathfrak{D}_{0,1/2})$ and the π-meson together with the θ-meson by an antisymmetric tensor in S_4.

Introduction

In order to interpret the notable stability of Λ-particles,[1] Pais[2] assumed the so-called " even-odd rule ". Recently he has proposed a new theory[3] of Λ-particles which gives a foundation to the even-odd rule. In this new theory it was supposed that the wave functions bescribing elementary particles, especially nucleons, Λ-particles and mesons, are functions not only of space-time variables x_μ but also of new variables concerning the ω-space. And the selection rule deduced from the parity of ω-space (ω-parity) corresponds to the even-odd rule. On the other hand Gell-Mann[4], and Nakano and Nishijima[5] obtained the even-odd rule from the charge independent character of strong interactions by assigning the integer isotopic spin and ordinary spin $1/2$ to Λ-particles.

As the interactions which produce Λ-and θ-particles by pion-nuclen or nucleon-nucleon reactions are rather strong, their effect will appear in the phenomena concerning the nuclear or pion-nucleon reactions. There seems, however, to be no empirical evidence of destroying the charge independent character of these strong interactions. So we may assume here the charge independence for the strong interactions between Λ-particles and nucleons.

In this paper we introduce the four dimensional charge space and get the representation of the full rotation group in the three-dimensional charge space from that of the four-dimensional full rotation group. The charge independence is guaranteed by requiring the invariance of the interaction Lagrangian under proper rotations in the three-dimensional charge space. Furthermore, requiring the invariance under the three-dimensional space inversion, we shall obtain the selection rule which corresponds to the even-odd rule. The conservation of nucleon family is derived from the requirement of the gauge invariance. This procedure to obtain the even-odd rule is equivalent to that of Pais' new theory, but

we do not use the ω-space explicity but only representation matrices.

A few years ago Utiyama had a chance of considering a generalization of ordinary τ-matrices from an academic point of view. However it was not intended to give any correspondence between consequences of the formal theory and experimental facts, because in that time one had naturally no information on the new particles. Accordingly any further development was not attempted and the theory remained unpublished.

Recently Nakano[6] has proposed a new theory of elementary particles which is quite similar to that of Utiyama though Nakano started from a different point of view, namely from that of a rigid-body-model of elementary particles.

In the present paper the out-line of our theory of A-particles will be given, the detailed discussion on its physical bearing will be published elsewhere.

§1. Mathematical consideration

The interaction Lagrangian density of a meson-nucleon system is given in the charge independent theory as

$$g\,\vec{\phi}\,\bar{\psi}\gamma_5\,\vec{\tau}\,\psi,$$

where $\vec{\tau}$ and $\vec{\phi}$ are defined as follows:

$$\tau_1 = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \tau_2 = \frac{1}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \tau_3 = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad (1\cdot1)$$

and

$$\phi_1 = \frac{1}{\sqrt{2}}\cdot(\phi+\phi^*), \quad \phi_2 = -\frac{i}{\sqrt{2}}\cdot(\phi-\phi^*),$$

$$\phi_3 = (\text{neutral meson field}).$$

The expression $(1\cdot1)$ is invariant under the proper rotation group G_3 in the 3-dimensional τ-space S_3 (Pais called S_3 "ω-space"). Explaining in more detail, we suppose that the vector $\vec{\phi}$ is transformed under an infinitesimal rotation as

$$\phi_i \to \phi_i' = \phi_i + \sum_{k=1}^{3}\varepsilon_{ik}\phi_k, \quad (i=1, 2, 3), \quad \varepsilon_{ik} = -\varepsilon_{ki}. \qquad (1\cdot2)$$

Together with $(1\cdot2)$, ψ is supposed to be transformed as

$$\psi \to \psi' = \{1+(i/2)\sum_{j,k=1}^{3}\varepsilon_{jk}\tau^{jk}\}\,\psi, \qquad (1\cdot3)$$

where we have put

$$\tau_1 = \tau^{23}, \quad \tau_2 = \tau^{31}, \quad \tau_3 = \tau^{12}.$$

It is obvious that $(1\cdot1)$ is invariant under these transformations.

In the above argument τ is considered to be an antisymmetric tensor rather than a vector in S_3.

Accordingly it is natural to interpret the vector $\vec{\phi}$ as an antisymmetric tensor ϕ_{ik} in

order to maintain the invariance of $(1 \cdot 1)$ under the full rotation group G_3' in S_3. Thus $(1 \cdot 1)$ will now be written as

$$(1/2)g\phi_{ik}\bar{\psi}\gamma_5\tau^{ik}\psi \tag{1.5}$$

(in what follows we shall omit summation-notation concerning suffices in τ-space).

However, if the wave function of meson is supposed to have a vector character under G_3', τ^{ik} must be replaced with a new matrix vector Γ'_i, in terms of which $(1 \cdot 1)$ is rewritten as

$$g'\phi_i\bar{\psi}\Lambda\Gamma'_i\psi . \tag{1.5}$$

Here Λ has been put equal to the unit matrix or γ_5 according to whether ϕ_i is a scalar or pseudoscalar in the ordinary space-time respectively.

Now in order to make $(1 \cdot 5)$ invariant under G_3', Γ_i must satisfy the following commutation relations :

$$[\tau^{mn}, \Gamma_i] = i\{\delta_{mi}\Gamma_n - \delta_{ni}\Gamma_m\} \tag{1.6}$$

and

$$\Gamma_i I + I\Gamma'_i = 0 , \tag{1.7}$$

where I is the inversion-operator by means of which ψ is transformed under the inversion in S_3 as follows :

$$\psi \to \psi' = I\psi$$

The matrices τ's have so far been considered as a particular representation of the generating operators of G_3, that is, as an irreducible representation $\mathfrak{D}_{1/2}$ of G_3. In other words, it has been assumed that the nucleon field ψ gives rise to the irreducible representation $\mathfrak{D}_{1/2}$ of G_3, and similarly the meson field corresponds to $\mathfrak{D}_1$.

Now let us propose that wave functions of every elementary particle are transformed according to an irreducible representation of G_3 under any rotation in S_3. That is to say, wave function ψ_α's $(\alpha = 1, 2, \cdots, 2\lambda+1, \ \lambda = 0, 1/2, 1, \cdots)^{*)}$ are to be transformed in the following way :

$$\psi_\alpha \to \psi_\alpha' = \sum_{\beta=1}^{2\lambda+1}\{\delta_{\alpha\beta} + (i/2)\varepsilon_{jk}(\alpha \mid D^{jk} \mid \beta)\}\psi_\beta \quad \text{(rotation)},$$

$$\tag{1.8}$$

$$\psi_\alpha \to \psi_\alpha' = \sum_{\beta=1}^{2\lambda+1}(\alpha \mid I \mid \beta)\psi_\beta \quad \text{(inversion in } S_3\text{)}.$$

Here the $(2\lambda+1)$-dimensional matrices D^{ik}'s make up an irreducible representation $\mathfrak{D}_\lambda$ of G_3.

According to this postulate, it is reasonable to replace $(1 \cdot 4)$ and $(1 \cdot 5)$ with the more general expressions

---

*) The suffix α refers to the τ-space. In this paper we omit all the spinor suffices concerning the ordinary space-time.

$$(1/2)g\phi_{jk}\bar{\psi}\gamma_5 D^{jk}\psi \qquad\qquad (1\cdot9)$$

and

$$g'\phi_j\bar{\psi}\Lambda\Gamma'_j\psi, \qquad\qquad (1\cdot10)$$

where Γ'_j's are now matrices of degree $(2\lambda+1)$ and satisfy the commutation relations

$$[D^{mn}, \Gamma'_j]=i\{\delta_{mj}\Gamma'_n-\delta_{nj}\Gamma'_m\}, \ \Gamma'_j I+I\Gamma'_j=0. \qquad (1\cdot11)$$

Now, since the inversion in S_3 is commutable with any rotation, I is equal to plus or minus unit matrix of degree $(2\lambda+1)$ in so far as D^{jk}'s are matrices of an irreducible representation (Schur-Burnside Theorem). In this case we have only a trivial solution of $(1\cdot11)$, i. e., $\Gamma'_j=0$. Therefore in order to deal with non-trivial cases, it is necessary to take into account a reducible representation of G_3.

Let us concern ourselves with any irreducible set of matrices D^{jk}, I and Γ'_j satisfying $(1\cdot11)$. In such a set D^{jk}'s alone are no longer irreducible and can be splitted into several irreducible representations of G_3 by a suitable transformation, and can be written as

$$(1\cdot12)$$

On account of the commutability between I and D^{jk}'s, I turns out in the above special representation as follows :

$$(1\cdot13)$$

where each E_j is a unit matrix of a particular degree d_j, and c_j is equal to $+1$ or -1. Furthermore ψ turns out to be splitted into several parts each one of which belongs respectively to a particular irreducible invariant sub-space of the representation-space of G_3.

Now we have at hands two alternative ways of getting concrete representations of D, I and Γ'.

The first way is to solve $(1\cdot11)$ without any assumption. The second one is to solve the problem by considering a four-dimensional Euclidian τ-space S_4 in which S_3 is imbedded as a subspace. These two lines of reasoning are quite similar to those which have been taken in Bhabha's paper.[7]

In the present paper we shall be concerned only with the second way whereas the study in the first one will be published in the near future.

§2. Four-dimensional treatment of τ-space

In order to solve $(1 \cdot 11)$, let us consider a four-dimensional Euclidian τ-space S_4, and denote the generating operators of the proper rotation-group G_4 in S_4 by $D^{\mu\nu}(\mu, \nu = 1, 2, \cdots, 4)$. These operators satisfy the following commutation-relations

$$[D^{mn}, D^{rs}] = i\{\delta_{mr}D^{ns} - \delta_{ms}D^{nr} - \delta_{nr}D^{ms} + \delta_{ns}D^{mn}\},$$ (2·1a)

$$(m, n, r, s = 1, 2, 3)$$

$$[D^{mn}, D^{j4}] = i\{\delta_{mj}D^{n4} - \delta_{nj}D^{m4}\}$$ (2·1b)

and

$$[D^{j4}, D^{k4}] = iD^{jk}.$$ (2·1c)

Let us further consider the 3-dimensional inversion I in S_4, by means of which any vector χ_μ in S_4 is transformed as

$$\chi_\mu \to \chi_\nu' = (-\chi_1, -\chi_2, -\chi_3, +\chi_4).^{*)}$$

I satisfies the following commutation-relations

$$[D^{mn}, I] = 0$$ (2·2a)

and

$$D^{m4}I + ID^{m4} = 0.$$ (2·2b)

By comparing $(2 \cdot 1b)$ and $(2 \cdot 2b)$ with the relations $(1 \cdot 11)$, it will be seen that we can identify Γ'_j with D^{j4} if we impose the further relations

$$[\Gamma_j, \Gamma_k] = iD^{jk}$$ (2·3)

on Γ'_j's beforehand.

In this point of view the meson field ϕ_i should be denoted by the four-dimensional notation ϕ_{i4}. Therefore the interaction Lagrangian density can be written in a four-dimensionally invariant form as

$$\sum_{\mu,\nu=1}^{4} (1/2) g\phi_{\mu\nu}\bar{\psi}\Lambda D^{\mu\nu}\psi,$$ (2·4)

provided that g' is taken to be equal to g.

From the view point mentioned above we can easily give concrete representations of D^{jk}, Γ_j and I. Irreducible representations of G_4 are quite familiar to us and they are denoted by $\mathfrak{D}_{j,j'}(j, j' = 0, 1/2, 1, \cdots)$. Furthermore any irreducible representation $R_{j+j', j-j'}$ of the full rotation group G_4' in S_4 is described as follows:

i) *Case of $j \neq j'$ (We may assume $j > j'$.)*

For any element $X \in G_4$, the representation $D(X)$ is given by

---

*) Here we use the same notations both for operators and for their representation matrices.

$$D(X) \equiv R_{j+j', j-j'}(X) = \left(\begin{array}{c|c} \mathfrak{D}_{j,j'} & 0 \\ \hline 0 & \mathfrak{D}_{j',j} \end{array} \right), \tag{2.5}$$

while our inversion I is represented by

$$I \equiv R_{j+j', j-j'}(I) = \left(\begin{array}{c|c} 0 & E \\ \hline E & 0 \end{array} \right), \tag{2.6}$$

where E is a unit matrix of degree $(2j+1)(2j'+1)$.

After a unitary transformation by

$$U = \frac{1}{\sqrt{2}} \left(\begin{array}{c|c} E & E \\ \hline E & -E \end{array} \right) = U^{-1}$$

(2.5) and (2.6) become

$$D(X) \rightarrow U^{-1}D(X)U = \left(\begin{array}{c|c} \dfrac{D_1 + D_2}{2} & \dfrac{D_1 - D_2}{2} \\ \hline \dfrac{D_1 - D_2}{2} & \dfrac{D_1 + D_2}{2} \end{array} \right) \tag{2.5'}$$

and

$$I \rightarrow U^{-1}IU = \left(\begin{array}{c|c} E & 0 \\ \hline 0 & -E \end{array} \right), \tag{2.6'}$$

where matrices of $\mathfrak{D}_{j,j'}$ and $\mathfrak{D}_{j',j}$ have been denoted by D_1 and D_2 respectively.

Now putting $X = 1 + i\varepsilon D^{mn}$ $(m, n = 1, 2, 3)$, we have

$$U^{-1}D(D^{mn})U \equiv D^{mn} = \left(\begin{array}{c|c} \diagup\!\!\!\!\diagup & 0 \\ \hline 0 & \diagup\!\!\!\!\diagup \end{array} \right), (m, n = 1, 2, 3), \tag{2.7}$$

where D^{mn} on the left-hand side of the above expression means the generating element of G_4, while that in the middle term stands for the corresponding representation matrix. The matrix form of D^{mn} on the right-hand side has its origin in the relations (2.2a).

On the other hand if we put $X = 1 + i\varepsilon D^{m4}$, we have

$$D^{m4} = \left(\begin{array}{c|c} 0 & \diagup\!\!\!\!\diagup \\ \hline \diagup\!\!\!\!\diagup & 0 \end{array} \right) \tag{2.8}$$

on account of $(2 \cdot 2b)$.

According to the well-known theorem of the group theory, $\mathfrak{D}_{j,j'}$ turns out to be a direct sum of several irreducible representations $\mathfrak{D}_k$ of G_3' when we are concerned only with the subgroup G_3' instead of the original group G_4', i. e.,[9]

$$\mathfrak{D}_{j,j'} = \mathfrak{D}_{j+j'} \dot{+} \mathfrak{D}^*_{j+j'-1} \dot{+} \mathfrak{D}_{j+j'-2} \dot{+} \cdots \dot{+} \begin{cases} \mathfrak{D}_{j-j'}, & j' = \text{integer} \\ \mathfrak{D}^*_{j-j'}, & j' = \text{half-odd-integer.} \end{cases} \qquad (2 \cdot 9)$$

Here $\mathfrak{D}_k^*$ means

$$\mathfrak{D}_k^* = \mathfrak{D}_k \times \mathfrak{D}_0^*$$

and $\mathfrak{D}_0^*$ is a one-dimensional representation of G_3' and is defined as follows:

$$\mathfrak{D}_0^*(X) = \begin{cases} +1: & \text{proper } X \epsilon G_3' \\ -1: & \text{improper } X \epsilon G_3'. \end{cases}$$

Now let us take a special representation of G_4' where D^{mn}'s $(m, n = 1, 2, 3)$ and I have the following forms:

$$D^{mn} = \begin{pmatrix} \mathfrak{D}_{j+j'} & & & & & & \\ & \mathfrak{D}_{j+j'-1} & & & & 0 & \\ & & \ddots & & & & \\ & & & \mathfrak{D}_{j-j'} & & & \\ \hline & & & & \mathfrak{D}^*_{j+j'} & & \\ & & & & & \mathfrak{D}^*_{j+j'-1} & \\ & 0 & & & & & \ddots \\ & & & & & & & \mathfrak{D}^*_{j-j'} \end{pmatrix}, \quad I = \left(\begin{array}{c|c} E & 0 \\ \hline 0 & -E \end{array} \right) \qquad (2 \cdot 10)$$

$$(m, n = 1, 2, 3)$$

This representation has just the same form with what was considered in $(1 \cdot 12)$. As stated in § 1, ψ_α's are now splitted into several parts, each one of which corresponds to an elementary particle with a definite τ-spin and parity for inversion in S_3 according to our view point stated on p. 413. Following Pais, let us call the parity of an elementary particle odd or even according to whether the wave function of this particle changes its sign or not under inversion. Then, from $(2 \cdot 8)$ it is easily seen that matrix elements of D^{j4}'s have non-vanishing values only when these elements connect two states (of elementary particles) with different parities.

ii) *Case of $j = j'$*

In this case the representation of G_4 which is obtained by the principle of selection from the irreducible representation $R_{2j,0}$ of G_4' is also irreducible.[10] For any proper element $X \epsilon G_4$, the representation $D(X)$ is given by

$$D(X) = R_{2j,0}(X) = \mathfrak{D}_{j,j}(X).$$

Now let us choose the special representation where D^{jk}'s $(j, k = 1, 2, 3)$ have a form of direct sum of several irreducible representations of G_3', i. e.,

$$D^{ik} = \begin{pmatrix} \mathcal{D}_{2j} & & & \\ & \mathcal{D}_{2j-1}^{*} & & \\ & & \mathcal{D}_{2j-2} & \\ & & & \ddots \\ & & & & \mathcal{D}_0 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} \mathcal{D}_{2j} & & & \\ & \mathcal{D}_{2j-1}^{*} & & \\ & & \mathcal{D}_{2j-2} & \\ & & & \ddots \\ & & & & \mathcal{D}_0^{*} \end{pmatrix} \cdot \quad (2\cdot 11)$$

$$(i,\, k = 1,\, 2,\, 3)$$

according to whether j is an integer or a half-odd-integer. In this representation I has the form

$$I = \begin{pmatrix} E_{4j+1} & & & \\ & -E_{4j-1} & & \\ & & E_{4j-3} & \\ & & & \ddots \\ & & & & +1 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} E_{4j+1} & & & \\ & -E_{4j-1} & & \\ & & E_{4j-3} & \\ & & & \ddots \\ & & & & -1 \end{pmatrix}, \quad (2\cdot 12)$$

where E_k means a unit matrix of degree k. Furthermore owing to the relation $(2\cdot 2b)$, $(2\cdot 1b)$ and $(2\cdot 1c)$, D^{i4}'s $(i=1, 2, 3)$ have the following form

$$D^{i4} = \begin{pmatrix} 0 & \diagdown & 0 \\ \diagdown & 0 & \diagdown \\ & \diagdown & 0 & \diagdown \\ 0 & & \diagdown & 0 \end{pmatrix} \qquad (i=1, 2, 3) \qquad (2\cdot 13)$$

Thus in this case, matrix elements of D^{i4}'s have non-vanishing values only when they connect two states with different parities and τ-spins.

Examples

i) $j=1/2,\ j'=0$

$$D^{23} = \begin{pmatrix} \tau^{23} & 0 \\ 0 & \tau^{23} \end{pmatrix}, \qquad D^{31} = \begin{pmatrix} \tau^{31} & 0 \\ 0 & \tau^{31} \end{pmatrix}, \qquad D^{12} = \begin{pmatrix} \tau^{12} & 0 \\ 0 & \tau^{12} \end{pmatrix},$$

$$D^{11} = \begin{pmatrix} 0 & \tau^{23} \\ \tau^{23} & 0 \end{pmatrix}, \qquad D^{21} = \begin{pmatrix} 0 & \tau^{31} \\ \tau^{31} & 0 \end{pmatrix}, \qquad D^{34} = \begin{pmatrix} 0 & \tau^{12} \\ \tau^{12} & 0 \end{pmatrix},$$

$$I = \begin{pmatrix} E_{\frac{1}{2}} & 0 \\ 0 & -E_{\frac{1}{2}} \end{pmatrix}.$$

ii) $j=j'=1/2$

$$D^{23}=\begin{pmatrix} 0 & \dfrac{1}{\sqrt{2}} & 0 & \\ \dfrac{1}{\sqrt{2}} & 0 & \dfrac{1}{\sqrt{2}} & 0 \\ 0 & \dfrac{1}{\sqrt{2}} & 0 & \\ \hline & 0 & & 0 \end{pmatrix}, \quad D^{31}=\begin{pmatrix} 0 & -\dfrac{i}{\sqrt{2}} & 0 & \\ \dfrac{i}{\sqrt{2}} & 0 & -\dfrac{i}{\sqrt{2}} & 0 \\ 0 & \dfrac{i}{\sqrt{2}} & 0 & \\ \hline & 0 & & 0 \end{pmatrix},$$

$$D^{12}=\begin{pmatrix} 1 & 0 & 0 & \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & \\ \hline & 0 & & 0 \end{pmatrix}, \quad D^{14}=\begin{pmatrix} & & & -\dfrac{1}{\sqrt{2}} \\ & 0 & & 0 \\ & & & \dfrac{1}{\sqrt{2}} \\ \hline -\dfrac{1}{\sqrt{2}} & 0 & \dfrac{1}{\sqrt{2}} & 0 \end{pmatrix},$$

$$D^{24}=\begin{pmatrix} & & & \dfrac{i}{\sqrt{2}} \\ & 0 & & 0 \\ & & & \dfrac{i}{\sqrt{2}} \\ \hline -\dfrac{i}{\sqrt{2}} & 0 & -\dfrac{i}{\sqrt{2}} & 0 \end{pmatrix}, \quad D^{34}=\begin{pmatrix} & & & 0 \\ & 0 & & 1 \\ & & & 0 \\ \hline 0 & 1 & 0 & 0 \end{pmatrix},$$

$$I=\begin{pmatrix} 1 & 0 & 0 & \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \\ \hline & 0 & & -1 \end{pmatrix}.$$

§3. Fundamental equations

The Lagrangian density of the meson-nucleon system in the current theory is written as

$$L=i\bar\psi\,(\gamma^\mu\partial_\mu+M)\,\psi-(1/2)\sum_{j=1}^{3}(\partial_\mu\phi_j\partial_\mu\phi_j+\mu^2\phi_j\phi_j)+g\sum_{j=1}^{3}\phi_j\bar\psi\gamma_5\tau_j\psi.$$

According to the view point so far advanced, it is natural to generalize the above expression in the following form :

$$L=i\bar\psi\,(\gamma^\mu\partial_\mu+\Sigma)\,\psi-(1/4)\sum_{j,k=1}^{4}\partial_\mu\phi_{jk}\partial_\mu\phi_{jk}$$

$$- (1/4) \sum_{j,k,l,m=1}^{4} \phi_{jk}(jk \mid \mu^2 \mid lm)\phi_{lm} \qquad (3\cdot1)$$

$$+ (1/4)g \sum_{j,k,l,m=1}^{4} \phi_{jk}\overline{\psi}(jk \mid \Lambda \mid lm)D^{lm}\psi.$$

Here Σ and μ^2 are mass-operators of ψ- and ϕ-field respectively, and are supposed to be defined as follows:

$$\Sigma = M_0 + (1/2)M_1 \sum_{i,k=1}^{3}(D^{ik})^2 + M_2 \sum_{j=1}^{3}(D^{j4})^2 + M_3 I + M_4 D^{12},$$

$$\mu^2 = \mu_0^2 + (1/2)\mu_1^2 \sum_{i,k=1}^{3}(D'^{ik})^2 + \mu_2^2 \sum_{j=1}^{3}(D'^{j4})^2 + \mu_3^2 I' + \mu_4^2 D'^{12}, \qquad (3\cdot2)$$

M_α and $\mu_\alpha^2 (\alpha=0, 1, \cdots, 4)$ are some constants to be determined later. The expression $(3\cdot2)$ is the simplest one being invariant under any transformation of G_3' (except the last term which gives rise to a small mass-difference between isobaric states with a fixed τ-spin) and has a diagonal form in the particular representation as stated in $(2\cdot11)$ and $(2\cdot12)$. In $(3\cdot2)$ the D'^{ik}'s are the representation matrices, in terms of which ϕ_{ik}'s are transformed under any rotation in S_4 as follows:

$$\phi_{ik} \rightarrow \phi_{ik}' = \phi_{ik} + \frac{1}{2} \sum_{r,s=1}^{4} \varepsilon_{rs} \sum_{m,n=1}^{4} (ik \mid D'^{rs} \mid mn)\phi_{mn} \quad (i, k=1\rightarrow4).$$

Similarly I' is a matrix which gives rise to a transformation of ϕ_{ik} corresponding to the inversion. The matrix Λ in $(3\cdot1)$ is to be determined according to what kinds of transformation characters the ϕ-fields have.

Since there is no experimental evidence of denying the existence of charged particles with two-π-meson decay-mode, it may be possible to assume that the particles of these types together with θ^0-particle correspond to three different isobaric states of a field of θ-particle with τ-spin 1. Accordingly let us describe this field by our field components ϕ_{i4}, whereas the ordinary meson field by $\phi_{ik}(i, k=1, 2, 3)$. Since the θ-meson field is tentatively interpreted to be a scalar (or vector) from its decay mode, Λ in $(3\cdot1)$ should be chosen as follows:

$$(ik \mid \Lambda \mid lm) = \text{antisymmetric with regard to interchange } (ik) \text{ and interchange } (lm)$$

$$= \begin{cases} (1/2)(\delta_{il}\delta_{km} - \delta_{kl}\delta_{im})\gamma_5 & (i, k, l, m=1, 2, 3) \\ \delta_{il} & (i, l=1, 2, 3, k=m=4). \end{cases}$$

Now let us choose a special representation $R_{j+j',j-j'}$ of G_4' where D^{ik} and I have the particular forms of $(2\cdot10)$ and further D^{12} is diagonal. In such a representation ψ becomes to be splitted into several parts as follow:

$$\psi = \begin{pmatrix} \psi^1 \\ \psi^2 \\ \vdots \end{pmatrix},$$

where each $\psi^k(k=1, 2, \cdots)$ spans a particular irreducible invariant sub-space for G_3'.

The interaction term having D^{ik}'s $(i, k = 1, 2, 3)$ as a factor gives rise to transitions within each subspace ψ^{Am} accompanied with emission or absorption of an ordinary π-meson, while the term having D^{i4}'s those between ψ^{Am} and ψ^{An} $(m \gtreqless n)$ with emission or absorption of a θ-meson.

In what follows we shall be concerned only with the special representation $R_{1/2, 1/2}$ $(j = 1/2, j' = 0)$, because the nucleon field customarily is considered to be that with τ-spin 1/2. In this case ψ has two invariant irreducible parts ψ^A and ψ^B, i. e.,

$$\psi = \begin{pmatrix} \psi^A \\ \psi^B \end{pmatrix}$$

each $\psi^i (i = A, B)$ has two components ψ^{i1} and ψ^{i2} respectively albeit they have the ordinary spinor components in their turn.

Now let us propose that ψ^{A1} and ψ^{A2} represent proton and neutron fields respectively, while ψ^{B1} and ψ^{B2} stand for Λ^+[11)] and Λ^0.

Under this assumption we have the following transition modes:

a)   the interaction term

$$(1/2) g \sum_{i,k=1}^{3} \phi_{ik} \bar{\psi} \gamma_5 D^{ik} \psi$$

gives rise to transitions

$$N \rightleftarrows N + \pi, \quad \Lambda \rightleftarrows \Lambda + \pi, \quad N + N \rightleftarrows \pi, \quad \Lambda + \Lambda \rightleftarrows \pi$$

b)   the interaction term

$$g \sum_{i=1}^{3} \phi_{i4} \bar{\psi} D^{i4} \psi$$

gives

$$N + \Lambda \rightleftarrows \theta, \quad N \rightleftarrows \Lambda + \theta, \quad \Lambda \rightleftarrows N + \theta,$$

where nucleons are denoted by N.

According to our terminology stated in § 2, nucleon and π-meson have the even parity, while Λ-particle and θ-meson the odd parity as easily seen from the first example of § 2. From a) and b) we see that the parities are conserved through transitions. This conservation of parity corresponds to Pais' "even-odd rule" and is nothing but the consequence of the invariance of the interaction under the inversion.

In order to take account of the electromagnetic interaccion, let us replace (3·1) with the following:

$$L = i\bar{\psi}[\gamma^\mu (\partial_\mu - ieCA_\mu(x)) + \Sigma]\psi$$
$$- (1/4) \sum_{i,k=1}^{4} \{\partial_\mu \phi_{ik} - ieA_\mu(x)(D'^{12}\phi)_{ik}\}^\dagger \cdot$$
$$\cdot \{\partial^\mu \phi_{ik} - ieA^\mu(x)(D'^{12}\phi)_{ik}\} \tag{3·3}$$
$$- (1/4) \sum_{i,k,l,m=1}^{4} \phi_{ik}(ik \mid \mu^2 \mid lm)\phi_{lm}$$

$$+ (1/4)\, g \sum_{i,k,l,m=1}^{4} \phi_{ik}\bar{\psi}(ik \mid \Lambda \mid lm)D_{lm}\psi + L_{el.mg.},$$

where C is given by

$$C = -\frac{1}{2}\begin{pmatrix} E_2 & 0 \\ 0 & E_2 \end{pmatrix} + D^{12} = \begin{pmatrix} 0 & 0 & & 0 \\ 0 & 0 & & \\ & & 1 & \bar{0} \\ 0 & & 0 & 0 \end{pmatrix}.$$

On account of the definition of D'^{12}, i. e.,

$$(jk \mid D'^{12} \mid lm) = i[\delta_{km}(\delta_{j1}\delta_{l2} - \delta_{l1}\delta_{j2}) + \delta_{jl}(\delta_{k1}\delta_{m2} - \delta_{m1}\delta_{k2})],$$

we see that the factor $\{\cdots\}^{\dagger}$ in the second term of L is hermitian and is equal to $\{\cdots\}$.

Now we can easily see the γ-stability of the $\varLambda$-particle, because the electromagnetic interaction-term has the diagonal matrix C as a factor. (3.3) is invariant under the following two transformations.

i) *Gauge transformation*

This is the transformation

$$\psi \to \psi + \delta\psi = \psi + ie\lambda(x)\, C\psi,$$
$$\bar{\psi} \to \bar{\psi} + \delta\bar{\psi} = \bar{\psi} - ie\lambda(x)\bar{\psi}C,$$
$$\phi_{ik} \to \phi + \delta\phi_{ik} = \sum_{l,m=1}^{4}(ik \mid V \mid lm)\phi_{lm}, \qquad (3.4)$$
$$I = 1 + ie\lambda D^{12\prime},$$
$$A_{\mu} \to A_{\mu} + \frac{\partial\lambda}{\partial x_{\mu}}.$$

The invariance of the first and second terms in (3.3) under (3.4) is obvious, and that of the third term can be proved by virtue of the following relations

$$(ik \mid D'^{12} \mid lm) = - (lm \mid D'^{12} \mid ik)$$

and

$$[D'^{12}, \mu^2] = 0.$$

The variation of the fourth term in the first order in λ is

$$\delta\{\Sigma\phi_{ik}\bar{\psi}(ik \mid \Lambda \mid lm)D^{lm}\psi\}$$
$$= ie\lambda\Sigma\{(D'^{12}\phi)_{ik}\bar{\psi}(ik \mid \Lambda \mid lm)D^{lm}\psi$$
$$- \phi_{ik}\delta\bar{\psi}(ik \mid \Lambda \mid lm)[D'^{12}, D^{lm}]\psi\}.$$

This is just the expression for variation caused by the infinitesimal rotation about the third axis in S_3. Thus this expression vanishes on account of the second invariant character.

ii) *Rotation about the third axis in τ-space*

$$\psi \to \psi + ie\theta D'^{12}\psi,$$
$$\bar{\psi} \to \bar{\psi} - ie\theta\bar{\psi}D^{12}, \qquad (3.5)$$
$$\phi_{ik} \to \phi_{ik} + ie\theta(D'^{12}\phi)_{ik}.$$

It is quite obvious that $(3\cdot3)$ is invariant under the transformation $(3\cdot5)$.

From these invariant character we get two constants of motion, i. e., from i) we get

$$\frac{\partial j^{\mu}}{\partial x^{\mu}}=0 \quad \text{and} \quad \int j^{0}(x)dx=Q=\text{const.}, \tag{3.6}$$

where j^{μ} is defined by

$$j^{\mu}=\bar{\psi}\gamma^{\mu}C\psi+(i/4)\sum_{i,k=1}^{4}[\{\partial^{\mu}\phi_{ik}-ieA^{\mu}(D'^{12}\phi)_{ik}\}(D'^{12}\phi)_{ik} \\ -(\phi D'^{12})_{ik}\{\partial^{\mu}\phi_{ik}-ieA^{\mu}(D'^{12}\phi)_{ik}\}]. \tag{3.7}$$

$(3\cdot6)$ shows the conservation of the electric charge. On the other hand, from ii) we have another law of conservation, i. e.,

$$\frac{\partial S^{\mu}}{\partial x^{\mu}}=0 \quad \text{and} \quad S=\int S^{0}(x)dx=\text{const.}, \tag{3.8}$$

where S^{μ} is

$$S^{\lambda}=\bar{\psi}\gamma^{\mu}D^{12}\psi+(i/4)\sum_{i,k=1}^{4}[\{\partial^{\mu}\phi_{ik}-ieA^{\mu}(D'^{12}\phi)_{ik}\}(D'^{12}\phi)_{ik} \\ -(\phi D'^{12})_{ik}\{\partial^{\mu}\phi_{ik}-ieA^{\mu}(D'^{12}\phi)_{ik}\}]. \tag{3.9}$$

Combining $(3\cdot6)$ with $(3\cdot8)$, we have the third conservation-law

$$\frac{\partial(\bar{\psi}\gamma^{\mu}\psi)}{\partial x^{\mu}}=0. \tag{3.10}$$

$(3\cdot10)$ means the conservation of numbers of particles belonging to the nucleon family.

Besides the interaction term of $(3\cdot1)$, it is necessary to introduce a new type of interaction term, in order to explain the decay process

$$\Lambda^{0}\rightarrow P+\pi^{-},$$

because this process breaks down the conservation of parity and cannot be derived from the interaction term of $(3\cdot1)$. If we also assume charge independence for this process, we may choose the following expression

$$(1/4)f\sum_{i,k=1}^{4}\phi_{ik}\bar{\psi}(ik\mid\Lambda\mid lm)\underline{D}^{lm}\psi, \tag{3.11}$$

as the interaction Lagrangian. In $(3\cdot11)$, $\underline{D}$ is the dual tensor of D, i. e.,

$$\underline{D}^{lm}=(1/2)\sum_{n,p=1}^{4}\varepsilon_{lm,np}D^{np}.$$

The interaction $(3\cdot11)$ is pseudoscalar because it changes its sign under the inversion, and gives rise to transitions of the following modes;

$$N\rightleftarrows N+\theta, \quad N\rightleftarrows\Lambda+\pi, \quad N+N\rightleftarrows\theta, \quad \Lambda+N\rightleftarrows\pi,$$
$$\Lambda\rightleftarrows\Lambda+\theta, \quad \Lambda\rightleftarrows N+\pi, \quad \Lambda+\Lambda\rightleftarrows\theta.$$

On account of the empirical fact that the Λ-particle has a long life-time, the coupling constant f should be chosen to be extremely small compared with g as in Pais' theory.

So far we have been concerned only with the special case of $R_{1/2,1/2}$. However, in order to describe nucleons by a field with τ-spin $1/2$, we should choose in general a particular representation $R_{j+j',j-j'}$ with $j = j' + 1/2$ because in decomposing this representation into a direct sum of several irreducible representations of G_3', $\mathfrak{D}_{1/2}$ certainly appears in this direct sum only in the above case, i. e.,

$$R_{2j'+1/2,1/2} = \mathfrak{D}_{2j'+1/2} \dot{+} \mathfrak{D}_{2j'-1/2} \dot{+} \cdots \dot{+} \mathfrak{D}_{1/2}$$
$$\dot{+} \mathfrak{C}^*_{2j'+1/2} \dot{+} \mathfrak{C}^*_{2j'-1/2} \dot{+} \cdots \dot{+} \mathfrak{C}_{1/2}.$$

Therefore the Λ-particle can never have an interger τ-spin and further Boson-field π and θ have always τ-spin 1 in our theory. If an experimental fact be discovered which does not agree with the above statement, our theory will be false in so far as the nucleon is supposed to have τ-spin $1/2$.

The authors wish to thank Prof. Hayakawa and Dr. K. Nishijima for their valuable discussions and also Prof. Pais for his kindness of sending us the preprints of his works.

References

1) We use the new nomenclature for new particles through this paper. Cf. Physics Today, Vol. 6, No. 12 (1953), 24.
2) A. Pais, Phys. Rev. **86** (1952), 663.
3) A. Pais, Prog. Theor. Phys. **10** (1953), 457.
4) M. Gell-Mann, Phys. Rev. **92** (1953), 833.
5) T. Nakano and K. Nishijima, Prog. Theor. Phys. **10** (1953), 581.
6) T. Nakano, " *On Quantization of Rigid Sphere*," lecture at the sectional meeting for the theory of elementary particles held at Yukawa Hall, Kyoto, on May 3rd, 1953. Unpublished.
7) H. J. Bhabha, Rev. Mod. Phys. **17** (1945), 200.
8) Murnaghan, *The theory of group representation*, p. 262.
9) Murnaghan, *i id* p. 276.
10) Murnaghan, *ibid* p. 264.
11) The positively charged Λ has been observed by the Pasadena group: C. M. York, R. B. Leighton and E. K. Bojornerud, Phys. Rev. **90** (1953), 167.

On the Magnetic Anisotropy Energy of FeF$_2$

Kazuo NIIRA and Takehiko OGUCHI[*)]

Physics Department, Tokyo Institute of Technology Oh-okayama, Tokyo

The magnetic anisotropy energy of FeF$_2$ deduced from the measurement of Stout and Matarrese on the magnetic susceptibility of the single crystal has a large value of $5.3\,(M/M_0)^2$cm^{-1} per Fe ion, where M, (M_0) is the magnetization, (the one at $0°K$), and this cannot be interpreted in terms of the magnetic dipole-dipole interaction between Fe ions in the crystal. The origin of the anisotropy energy can be explained by the spin-orbit interaction in the Fe^{++} ion. We have calculated the second order perturbation energy due to this interaction, taking account of the crystalline Stark effect of the ground state of the Fe^{++} ion. The crystalline field can be determined by the anisotropy of the g-factor deduced from the experiment of Stout and Matarrese. The calculated anisotropy energy per ion due to the spin-orbit interaction is $7.6\,(M/M_0)^2$ cm^{-1}, in good agreement with the observed value.

§ 1.   Introduction

A recent measurement on the magnetic susceptibility of the single crystals of MnF$_2$, FeF$_2$ and CoF$_2$ by Stout and Matarrese[1)] shows a magnetic anisotropy in these antiferromagnetic substances. On the other hand, the anisotropy of the antiferromagnetic substance CuCl$_2\cdot$2H$_2$O was calculated by Moriya and Yosida[2)]. As mentioned in their paper, the origin of the anisotropy energy can be interpreted by one or more of the following terms: 1) the magnetic interaction between dipoles or multipoles in the crystal, 2) the anisotropic exchange interaction between ions, and 3) the spin-orbit and spin-spin interactions in the ion under the influence of the anisotropic crystalline field. The origin of the anisotropy energy of MnF$_2$ was interpreted by Keffer[3)] in terms of 1). The crystal structures of MnF$_2$, FeF$_2$ and CoF$_2$ are all of the rutile type of nearly equal lattice constants, while the anisotropy energy of FeF$_2$ or CoF$_2$ is nearly 100 times larger than that of MnF$_2$. Then, it is evident that the origin of it for FeF$_2$ or CoF$_2$ cannot be explained in terms of 1), but its observed large value can only be interpreted in terms of 3) for the case of FeF$_2$, as will be shown in the following sections.

§ 2.   Analysis of the observed data of Stout and Matarrese

The crystal structure of FeF$_2$ and the spin alignment of Fe^{++} ions at the ordered phase are shown in Fig. 1[4)].

[*)]   Present adress, Institute of Physics, Tokyo University of Education.

[a] site   [b] site
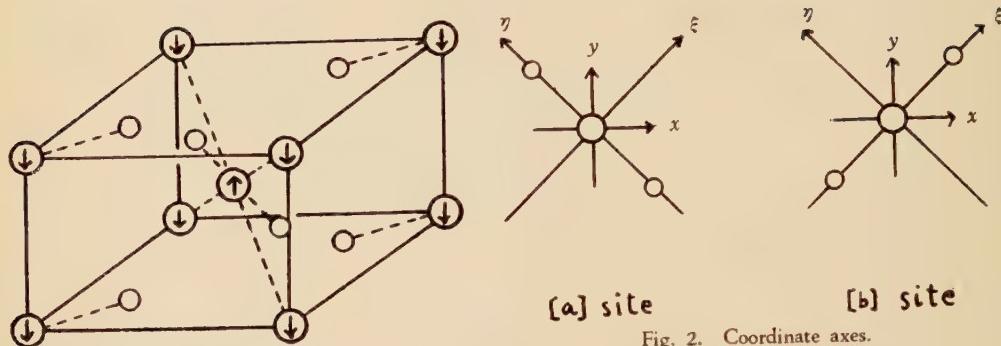
Fig. 2.   Coordinate axes.

○ Fe^{++}

○ F^-

Fig. 1.   Magnetic structure of FeF_2 showing the order and orientation of the Fe^{++} magnets.   The small circles correspond to fluorine sites.

We choose the coordinate axes x, y and z along the crystal axes, and ξ, η and ζ axes as shown in Fig. 2. Antiferromagnetic exchange force between the nearest Fe^{++} pair, the one on the body-center and the otherest corner lattice points (the distance between them being 3.25A,) will be dominant as compared to the exchange forces between the other pairs. As a matter of course, we have a spin alignment as shown in Fig. 1, which is the typical two-sublattice model of antiferromagnetism. We will indicate each of these two sublattices by $[a]$ or $[b]$ (Fig. 2).
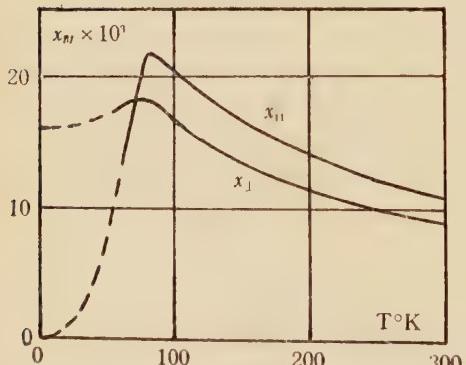
Fig. 3.   Molal magnetic susceptibility of FeF_2 parallel and perpendicular to ζ axis of crystal.
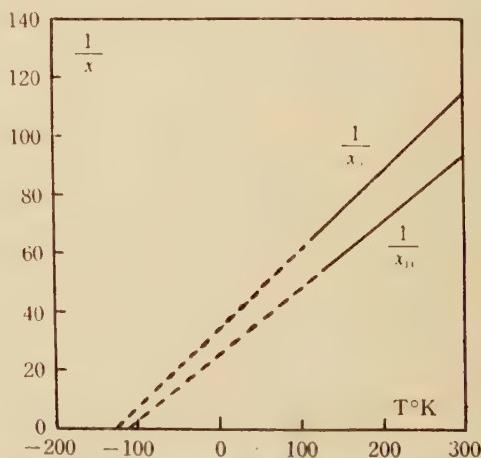
Fig. 4.   Reciprocal of molal magnetic susceptibility of FeF_2.

The observed value of the magnetic susceptibility of the single crystal FeF_2 by Stout and Matarrese[1] is shown in Fig. 3. The ordinate is the molal susceptibility, and $\chi_\parallel$ ($\chi_\perp$)

means the susceptibility along ζ (ξ or η) axis. The plots of $1/\chi_{\shortparallel}$ and $1/\chi_{\perp}$ against temperature are well manifested by the straight lines in the temperature region higher than about 100°K (Fig. 4).

Their temperature coefficients and the temperature $-T_0$ at which $1/\chi = 0$ are given in Table I.

Table I. Temperature coefficient of $1/\chi$ and T_0.

	temp. coef. (1/deg.)	T_0(°K)
$1/\chi_{\shortparallel}$	0.226	114.6
$1/\chi_{\perp}$	0.270	124.0

Considering from the crystal symmetry, the Lande's g-factors (tensors) for the $[a]$ and $[b]$ sites are given by

$$g^{(a)} = \begin{pmatrix} g_\xi & 0 & 0 \\ 0 & g_\eta & 0 \\ 0 & 0 & g_\zeta \end{pmatrix}, \quad g^{(b)} = \begin{pmatrix} g_\eta & 0 & 0 \\ 0 & g_\xi & 0 \\ 0 & 0 & g_\zeta \end{pmatrix}.$$

$$(2 \cdot 1)$$

When the temperature is higher than the Néel temperature of this antiferromagnetic substance, the introduction of the external magnetic field H_0 in ξ, η and ζ directions will give displacement δS_ξ, δS_η and δS_ζ to the every spin, respectively. Then, the effective field at $[a]$ or $[b]$ sites can be written respectively, using the molecular field approximation, as follows ;

$$H_\xi^{(a)} = H_0 - (2|J|z + D_\xi)\delta S_\xi / g_\xi \mu, \qquad (2 \cdot 2\xi)$$

$$H_\eta^{(a)} = H_0 - (2|J|z + D_\eta)\delta S_\eta / g_\eta \mu, \qquad (2 \cdot 2\eta)$$

$$H_\zeta^{(a)} = H_0 - (2|J|z + D_\zeta)\delta S_\zeta / g_\zeta \mu, \qquad (2 \cdot 2\zeta)$$

$$H_\xi^{(b)} = H_0 - (2|J|z + D_\eta)\delta S_\xi / g_\eta \mu, \qquad (2 \cdot 3\xi)$$

$$H_\eta^{(b)} = H_0 - (2|J|z + D_\xi)\delta S_\eta / g_\xi \mu, \qquad (2 \cdot 3\eta)$$

$$H_\zeta^{(b)} = H_0 - (2|J|z + D_\zeta)\delta S_\zeta / g_\zeta \mu, \qquad (2 \cdot 3\zeta)$$

where J means the exchange integral between the nearest neighbor Fe ions, z the number of the nearest neighbors ($z = 8$), μ the Bohr magneton. D_ξ, D_η and D_ζ are the coefficients of effective anisotropy field at $[a]$ site when the external field is applied in the ξ, η and ζ directions respectively. With (2.2ζ), (2.3ζ) the molar parallel susceptibility $\chi_{\shortparallel}$ is given by

$$\chi_{\shortparallel} = \chi_\zeta = \frac{Ng_{\shortparallel}^2 \mu^2 S(S+1)}{3kT + S(S+1)(2|J|z + D_{\shortparallel})}, \qquad (2 \cdot 4)$$

where

$$D_{\shortparallel} = D_\zeta, \quad g_{\shortparallel} = g_\zeta, \qquad (2 \cdot 5)$$

and N is Avogadro's number, and S is the spin quantum number for Fe ion. Similarly, from (2.2ξ), (2.3ξ) or from (2.2η), (2.3η), we obtain the molar perpendicular susceptibility $\chi_{\perp}$ as follows,

$$\chi_{\perp} = \chi_\xi = \chi_\eta = \frac{Ng_{\perp}^2 \mu^2 S(S+1)}{3kT + S(S+1)(2|J|z + D_{\perp})}, \qquad (2 \cdot 6)$$

where

$$D_\perp = \frac{D_\xi g_\xi + D_\eta g_\eta}{g_\xi + g_\eta}, \qquad g^2_\perp = 1/2(g_\xi^2 + g_\eta^2). \qquad (2 \cdot 7)$$

From (2.4) and (2.6),

the coef. of T in $1/\chi_\parallel$ is $3k/Ng_\parallel^2\mu^2 S(S+1)$,   (2·8a)

the coef. of T in $1/\chi_\perp$ is $3k/Ng_\perp^2\mu^2 S(S+1)$,   (2·8b)

and $g_\parallel$ and $g_\perp$ can be determined by comparing them with the observed data of Stout and Matarrese. Taking $S=2$ for Fe ion, we obtain from Table I and (2.8a,b)

$$g_\parallel = 2.45, \qquad\qquad (2 \cdot 9a)$$

$$g_\perp = 2.24. \qquad\qquad (2 \cdot 9b)$$

The mean value of g, i.e. $g_m = (3/1)g_\parallel + (3/2)g_\perp = 2.31$, is consistent with the observed value 2.32 from the neutron diffraction by Erickson[4]. T_0 of $1/\chi_\parallel$ and of $1/\chi_\perp$ are obtained by the comparison between Table I and (2.4) or (2.6):

$$\frac{S(S+1)}{3k}(2|J|z + D_\parallel) = 114.6 \ ^\circ K, \qquad (2.10a)$$

$$\frac{S(S+1)}{3k}(2|J|z + D_\perp) = 124.0 \ ^\circ K, \qquad (2.10b)$$

respectively. According to Van Vleck[5], the Néel temperature T_N is, if the anisotropy field were absent, given by

$$T_N = \frac{S(S+1)}{3k} 2|J|z, \qquad\qquad (2 \cdot 11)$$

whereas the observed T_N is 79°K. Then, we have

$$2|J|z = 39.5 \ k, \qquad\qquad (2 \cdot 12)$$

where k is the Boltzmann constant. On the other hand, when the anisotropy field is considerably large as in the present case, or the exchange interaction between non-neighbor ions is important, T_N should be modified from (2.11)[6]. But the final numerical value of our calculation is found to be unaffected by the choice between (2.12) and 55 k, the value inclusive of the anisotropy field, for the value of $2|J|z$. Therefore, we choose the former. Then, from (2.10a,b) and (2.12),

$$D_\parallel = 12.3k, \qquad\qquad (2 \cdot 13a)$$

$$D_\perp = 17.0k. \qquad\qquad (2 \cdot 13b)$$

By the consideration of the crystal symmetry the anisotropy energy of FeF_2 per mol can be written as

$$E_A = K_1(\alpha^2 + \beta^2) + K_2 \gamma^2 = N\{\kappa_1(\alpha^2 + \beta^2) + \kappa_2 \gamma^2\}, \qquad (2 \cdot 14)$$

where α, β and γ are the direction cosines of the spin axis with respect to the ξ, η and ζ axes, respectively.

When the external field H_0 is applied in ζ or ξ direction, the magnetization per mole is given by

$$M_{||} = Ng_{||}\mu \ \delta S_{||}, \tag{2.15a}$$

$$M_{\perp} = Ng_{\perp}\mu \ \delta S_{\perp}, \tag{2.15b}$$

respectively, and from (2.4) and (2.6), the effective fields are given by following forms:

$$H_{||} = H_0 - (2|J|z + D_{||})\frac{\partial M_{||}}{N(g_{||}\mu)^2}, \tag{2.16a}$$

$$H_{\perp} = H_0 - (2|J|z + D_{\perp})\frac{\partial M_{\perp}}{N(g_{\perp}\mu)^2}. \tag{2.16b}$$

Then we have

$$\kappa_1 = -M/N \cdot H_0 + 1/2 \cdot (2|J|z + D_{\perp})(M/Ng_{\perp}\mu)^2, \tag{2.17a}$$

$$\kappa_2 = -M/N \cdot H_0 + 1/2 \cdot (2|J|z + D_{||})(M/Ng_{||}\mu)^2. \tag{2.17b}$$

From (2.9), (2.12), (2.13) and (2.17), we can obtain the anisotropy energy per ion as

$$\kappa_1 - \kappa_2 = \frac{1}{2}(2|J|z)(1/g_{\perp}^2 - 1/g_{||}^2)\frac{M^2}{N^2\mu^2} + \frac{1}{2}\left(\frac{D_{\perp}}{g_{\perp}^2} - \frac{D_{||}}{g_{||}^2}\right)\frac{M^2}{N^2\mu^2}$$

$$= 3.2 \times 10^{-8} \ M^2 \ \mathrm{cm}^{-1}. \tag{2.18}$$

Making use of the magnetization of the sublattice at $0°K$:

$$M_0 = N/2 \cdot g_m \mu \ S, \tag{2.19}$$

we can rewrite (2.18) as

$$\kappa_1 - \kappa_2 = 5.3 \ (M/M_0)^2 \ \mathrm{cm}^{-1}. \tag{2.20}$$

This is the anisotropy energy of FeF$_2$ deduced from the susceptibilities observed by Stout and Matarrese.

§ 3. Electronic Eigenfunctions of Fe ion

We consider an Fe ion on a sublattice point in the crystal of FeF$_2$. The crystalline electric field potential V on the ion exerted by the surrounding ions can be expressed from the symmetry consideration as

$$V = A(2z^2 - x^2 - y^2) + Bxy, \tag{3.1}$$

where terms of the fourth degree are assumed negligible. Transforming the coordinate axes to the ξ, η and ζ axes shown in Fig. 2, we have for $[a]$ and $[b]$ sites,

$$V = (-A + B/2)\xi^2 + (-A - B/2)\eta^2 + 2A\zeta^2, \tag{3.2a}$$

and

$$V=(-A-B/2)\xi^2+(-A+B/2)\eta^2+2A\zeta^2, \tag{3·2b}$$

respectively. First, we will give eigenfunctions of Fe ion. Assuming the Hund rule for the free Fe ion, we find ^5D as the ground state of the ion. We assume that the ^5D remains the ground state even in the crystal, and that other states than it are of high energies and can be neglected in our calculation. The orbital part of the ^5D state is five-fold degenerated in the absence of the crystalline field, and each substate can be chosen as the following real orthonormal functions (we omit the radial function for brevity):

$$\begin{aligned}
\phi_2^+ &= 1/\sqrt{2} & (\phi_2+\phi_{-2}) \\
\phi_2^- &= 1/\sqrt{2}i & (\phi_2-\phi_{-2})
\end{aligned} \bigg\} = \sqrt{15/16\pi}\,\sin^2\theta \begin{cases} \cos 2\varphi \sim \xi^2-\eta^2, \\ \sin 2\varphi \sim \xi\eta, \end{cases}$$

$$\begin{aligned}
\phi_1^+ &= 1/\sqrt{2} & (\phi_1+\phi_{-1}) \\
\phi_1^- &= 1/\sqrt{2}i & (\phi_1-\phi_{-1})
\end{aligned} \bigg\} = \sqrt{15/16\pi}\,\sin 2\theta \begin{cases} \cos\varphi \sim \xi\zeta, \\ \sin\varphi \sim \eta\zeta, \end{cases} \tag{3·3}$$

$$\phi_0 \qquad\qquad\qquad = \sqrt{15/16\pi}\,(2-3\sin^2\theta) \sim 2\zeta^2-\xi^2-\eta^2,$$

where ϕ_m represents the spherical harmonics of the second degree, $P_2^{|m|}(\cos\theta)\,e^{im\varphi}$, and θ and φ are polar angles taking ζ axis as the pole. With respect to the basis (3.3) the crystalline potential (3.2a) is represented by the matrix:

$$\begin{array}{c|ccccc}
 & \phi_0 & \phi_2^+ & \phi_2^- & \phi_1^+ & \phi_1^- \\
\hline
\phi_0 & 4/7\langle A\rangle & 2\sqrt{3}/21\langle B\rangle & 0 & 0 & 0 \\
\phi_2^+ & 2\sqrt{3}/21\langle B\rangle & -4/7\langle A\rangle & 0 & 0 & 0 \\
\phi_2^- & 0 & 0 & -4/7\langle A\rangle & 0 & 0 \\
\phi_1^+ & 0 & 0 & 0 & 2/7\langle A\rangle+1/7\langle B\rangle & 0 \\
\phi_1^- & 0 & 0 & 0 & 0 & 2/7\langle A\rangle-1/7\langle B\rangle
\end{array} \tag{3·4}$$

where $\langle A\rangle$ and $\langle B\rangle$ are the abbreviations for $A\langle r^2\rangle_{Av}$ and $B\langle r^2\rangle_{Av}$, respectively. From (3.4) the eigenfunctions and the energy values of the Fe ion in the crystal are immediately found to

$$\begin{aligned}
\psi_g &= c_0\phi_0 + c_2\phi_2', & E_g &= -4/7\sqrt{\langle A\rangle^2+1/12\langle B\rangle^2}, \\
\psi_h &= c_0'\phi_0 + c_2'\phi_2^+, & E_h &= 4/7\sqrt{\langle A\rangle^2+1/12\langle B\rangle^2}, \\
\psi_1 &= \phi_2^-, & E_1 &= -4/7\langle A\rangle, \\
\psi_2 &= \phi_1^+, & E_2 &= 2/7\langle A\rangle+1/7\langle B\rangle, \\
\psi_3 &= \phi_1^-, & E_3 &= 2/7\langle A\rangle-1/7\langle B\rangle,
\end{aligned} \tag{3·5}$$

where

$$\frac{c_0}{c_2} = \frac{2\sqrt{3}}{\langle B\rangle}(-\langle A\rangle+\sqrt{\langle A\rangle^2+1/12\langle B\rangle^2}), \qquad c_0^2+c_2^2=1,$$

$$\frac{c_0'}{c_2'} = \frac{2\sqrt{3}}{\langle B\rangle}(-\langle A\rangle-\sqrt{\langle A\rangle^2+1/12\langle B\rangle^2}), \qquad c_0'^2+c_2'^2=1. \tag{3·6}$$

It is clear from (3.5) that the lowest eigenvalue is always given by E_g, irrespective of the values of $\langle A \rangle$ and $\langle B \rangle$. The energy levels of Fe^{++} ion in the crystal are shown in Fig. 5, where we put for $\langle A \rangle$ and $\langle B \rangle$ the values given later (see § 4).

§ 4. Calculation of the anisotropy energy

In this section we will calculate the g-factor and the anisotropy energy, treating the spin-orbit interaction $\lambda(LS)$ and the spin-spin interaction in Fe ion as a perturbation. The method of our calculation follows nearly the same way as in the work by Moriya and Yosida[2], but, as the magnitude of the spin of Fe ion is 2, the "one atom model" is suitably applicable for the calculation of the anisotropy energy. As the result of it the second order perturbation of the spin-orbit interaction gives the major part of the anisotropy energy, the physical meaning of which can be understood quite clearly (see § 5, 7).

Fig. 5. Energy levels of Fe^{++} in the crystal of FeF_2.

Each state of the five wave functions (3.5) still has a five-fold degeneracy with respect to the spin state. By the first order perturbation of $\lambda(LS)^{*)}$ the ground state (ψ_g) splits into the following five states :

$$
\begin{aligned}
f_{g,2} &= \psi_g\,\theta_{2,2} - 4ic_2\,\gamma_1\psi_1\theta_{2,2} + (\sqrt{3}\,c_0 - c_2)\gamma_2\psi_2\theta_{2,1} + i(\sqrt{3}\,c_0 + c_2)\gamma_3\psi_3\theta_{2,1}, \\
f_{g,1} &= \psi_g\theta_{2,1} - 2ic_2\gamma_1\psi_1\theta_{2,1} + (\sqrt{3}\,c_0 - c_2)(-\theta_{2,2} + \sqrt{6}/2\,\theta_{2,0})\gamma_2\psi_2 \\
&\quad + i(\sqrt{3}\,c_0 + c_2)(\theta_{2,2} + \sqrt{6}/2\,\theta_{2,0})\gamma_3\psi_3, \\
f_{g,0} &= \psi_g\theta_{2,0} + \sqrt{6}/2(\sqrt{3}\,c_0 - c_2)(-\theta_{2,1} + \theta_{2,-1})\gamma_2\psi_2 + i\sqrt{6}/2(\sqrt{3}\,c_0 + c_2) \\
&\quad\qquad\qquad \cdot (\theta_{2,1} + \theta_{2,-1})\gamma_3\,\psi_3, \\
f_{g,-1} &= \psi_g\theta_{2,-1} + 2ic_2\gamma_1\psi_1\theta_{2,-1} + (\sqrt{3}\,c_0 - c_2)(-\sqrt{6}/2\,\theta_{2,0} + \theta_{2,-2})\gamma_2\,\psi_2 \\
&\quad + i(\sqrt{3}\,c_0 + c_2)(\sqrt{6}/2\,\theta_{2,0} + \theta_{2,-2})\gamma_3\psi_3, \\
f_{g,-2} &= \psi_g\,\theta_{-2,-2} + 4ic_2\gamma_1\psi_1\theta_{-2,-2} - (\sqrt{3}\,c_0 - c_2)\gamma_2\psi_2\theta_{2,-1} + i(\sqrt{3}\,c_0 + c_2) \\
&\quad\qquad\qquad \cdot\gamma_3\psi_3\theta_{2,-1},
\end{aligned}
\qquad (4 \cdot 1)
$$

where

$$
\gamma_k = \frac{\lambda}{E_k - E_g} \qquad (k = 1,\ 2,\ 3), \qquad\qquad (4.2)
$$

---

*) The spin-spin interaction can be neglected.

and $\Theta_{2,m}$ is the spin function and m is the quantum number of the ζ component of the spin ($m = -2,\cdots,2$). When we can neglect energetically higher states other than (4.1), the paramagnetic susceptibility above the critical temperature is expressed by[?]

$$\chi_i = N\mu^2 \sum_{m,m'} \frac{|(L_i + 2S_i)_{mm'}|^2}{5kT}, \tag{4.3}$$

where terms independent of temperature are omitted, and i means ξ, η and ζ. $(L_i + 2S_i)_{mm'}$'s are matrix elements of $L_i + 2S_i$ with respect to the bases (4.1). And their non-vanishing terms are:

$$(f_{g,2}|L_\xi + 2S_\xi|f_{g,1}) = (f_{g,-1}|L_\xi + 2S_\xi|f_{g,-2}) = 2 - 2(\sqrt{3}\,c_0 + c_2)^2\gamma_3,$$
$$(f_{g,1}|L_\xi + 2S_\xi|f_{g,0}) = (f_{g,0}|L_\xi + 2S_\xi|f_{g,-1}) = \sqrt{6} - \sqrt{6}\,(\sqrt{3}\,c_0 + c_2)^2\gamma_3,$$
$$(f_{g,2}|L_\eta + 2S_\eta|f_{g,1}) = (f_{g,-1}|L_\eta + 2S_\eta|f_{g,-2}) = i\{-2 + 2(\sqrt{3}\,c_0 - c_2)^2\gamma_2\},$$
$$(f_{g,1}|L_\eta + 2S_\eta|f_{g,0}) = (f_{g,0}|L_\eta + 2S_\eta|f_{g,-1}) = i\{-\sqrt{6} + \sqrt{6}\,(\sqrt{3}\,c_0 - c_2)^2\gamma_2\},$$
$$(f_{g,2}|L_\zeta + 2S_\zeta|f_{g,2}) = (f_{g,-2}|L_\zeta + 2S_\zeta|f_{g,-2}) = -4 + 16c_2^2\gamma_1, \tag{4.4}$$
$$(f_{g,1}|L_\zeta + 2S_\zeta|f_{g,1}) = (f_{g,-1}|L_\zeta + 2S_\zeta|f_{g,-1}) = -2 + 8c_2^2\gamma_1,$$

where terms higher than the second order in γ_i are neglected, for γ_i will be found small (see Eq. (4.15)). From (4.3) and (4.4)

$$\chi_\perp = 1/2 \cdot (\chi_\xi + \chi_\eta) = 2N\mu^2/kT \cdot \{4 - 4(\sqrt{3}\,c_0 + c_2)^2\gamma_3 - 4(\sqrt{3}\,c_0 - c_2)^2\gamma_2\}, \tag{4.5a}$$
$$\chi_{||} = \chi_\zeta = 2N\mu^2/kT \cdot \{4 - 32\ c_2^2\gamma_1\}. \tag{4.5b}$$

Or, remembering the spin of 2 for Fe ion, we can rewrite them with the g-factors:

$$\chi_\perp = 2N\mu^2/kT \cdot g_\perp^2, \tag{4.6a}$$
$$\chi_{||} = 2N\mu^2/kT \cdot g_{||}^2. \tag{4.6b}$$

Making use of the experimental values for the g-factors (2.9 a, b), we have from (4.5a,b) and (4.6 a, b),

$$(\sqrt{3}\,c_0 + c_2)^2\gamma_2 + (\sqrt{3}\,c_0 + c_2)^2\gamma_3 = -0.259, \tag{4.7}$$
$$c_2^2\,\gamma_1 \qquad\quad = -0.060. \tag{4.8}$$

From (4.8), (4.2) and (3.5)

$$\frac{\lambda}{\langle B \rangle} = -\frac{0.24}{7c_2^2}\{-p + \sqrt{p^2 + 1/12}\ \}, \tag{4.9}$$

where

$$p = \langle A \rangle / \langle B \rangle. \tag{4.10}$$

Inserting (4.9) into (4.7) and remembering (3.5, 6) for c_0, c_2, γ_2 and γ_3, we have

$$\left[\frac{\{1 + 2p + 2\sqrt{p^2 + 1/12}\}^2}{1 + 2p + 4\sqrt{p^2 + 1/12}} - \frac{\{1 - 2p - 2\sqrt{p^2 + 1/12}\}^2}{1 - 2p - 4\sqrt{p^2 + 1/12}} \right] (-p + \sqrt{p^2 + 1/12}))^3 = 0.04. \tag{4.11}$$

This equation has such a solution as

$$p = 0.03. \tag{4.12}$$

Then, we have

$$\langle A \rangle = -2.3\lambda, \tag{4.13}$$

$$\langle B \rangle = -77\lambda. \tag{4.14}$$

With these values we obtain

$$\gamma_1 = -0.088,$$
$$\gamma_2 = -0.041, \tag{4.15}$$
$$\gamma_3 = -0.417.$$

If we take the numerical value of λ for Fe^{++}[8], as -100 cm^{-1}, we have for the energy of each state in (3.5) as follows :

$$E_1 - E_g = 1140 \text{ cm}^{-1},$$
$$E_2 - E_g = 2440, \tag{4.16}$$
$$E_3 - E_g = 240,$$
$$E_h - E_g = 2540.$$

According to (4.15), γ_1 and γ_2 are small enough to allow us to neglect their squares, but γ_3 is not so small that the neglect of its square term would introduce some error in our result.

As even the energy difference between the ground state (ψ_g) and the lowest excited state (ψ_3) amounts to -2.4λ, which corresponds to $340°$K $(\lambda = -100$ cm$^{-1})$, we can safely assume that Fe ion is in the ground state at lower temperatures. Referring to the ground state we can calculate the perturbed energies due to the spin-orbit interaction $\lambda(\boldsymbol{LS})$ and due to the spin-spin interaction H_{ss} in the Fe ion. According to Pryce[9], H_{ss} can be expressed as

$$H_{ss} = -\rho\{1/2(\boldsymbol{L},\boldsymbol{S}) + (\boldsymbol{L},\boldsymbol{S})^2 - 1/3 \cdot L(L+1)S(S+1)\}, \tag{4.17}$$

where ρ is about 1cm^{-1} for Fe^{++} (^5D). The first order perturbation of H_{ss} gives rise the anisotropy energy ;

$$\Delta_1 E = -\rho\{c_2^2 S_\zeta^2 + (3c_0^2 + c_2^2 + 2\sqrt{3}\,c_0 c_2)S_\xi^2 + (3c_0^2 + c_2^2 - 2\sqrt{3}\,c_0 c_2)S_\eta^2\}, \tag{4.18}$$

while the first order perturbation of $\lambda(\boldsymbol{LS})$ vanishes. In the above expression we have left the spins in the operator form. The second order perturbation energy of $\lambda(\boldsymbol{LS})$ can be evaluated as

$$\Delta_2 E = -\lambda[4c_2^2\gamma_1 S_\zeta^2 + 1/2\{(\sqrt{3}\,c_0 - c_2)^2\gamma_2 + (\sqrt{3}\,c_0 + c_2)^2\gamma_3\}(S_\xi^2 + S_\eta^2)]. \tag{4.19}$$

As ρ is much smaller than $|\lambda|$, the second order perturbation of H_{ss} need not be considered. S_ξ, S_η, and S_ζ, which are operators, may be put equal to their expectation values :

$$S_\xi = S_\eta = \frac{M}{Ng_\perp\mu}, \qquad S_\zeta = \frac{M}{Ng_{||}\mu}, \tag{4.20}$$

where M is the magnetization per mol. Then, the anisotropy energy due to the spin-spin interaction is expressed as

$$(\kappa_1-\kappa_2)_{ss}=-\frac{\rho}{N^2\mu^2}\left(-\frac{4c_2^2}{g_\parallel^2}+\frac{3c_0^2+c_2^2}{g_\perp^2}\right)M^2, \qquad (4\cdot21)$$

and the corresponding quantity due to the spin-orbit interaction is

$$(\kappa_1-\kappa_2)_{so}=\frac{\lambda}{N^2\mu^2}\left[\frac{4c_2^2\gamma_1}{g_\parallel^2}-\frac{1}{2}\frac{\{(\sqrt{3}c_0-c_2)^2\gamma_2+(\sqrt{3}c_0+c_2)^2\gamma_3\}}{g_\perp^2}\right]M^2,$$

$$c_0^2=0.447, \quad c_2^2=0.553. \qquad (4\cdot22)$$

Inserting the values of $g_\parallel$ and $g_\perp$ and making use of (4.7) and (4.8), we find $-0.03\times 10^{-8} M^2$ cm^{-1} and $4.5\times10^{-8} M^2$ cm^{-1} for the right hand sides of (4.21) and (4.22), respectively. We can neglect the anisotropy energy due to the spin-spin interaction and we have $4.5\times10^{-8} M^2$ cm^{-1} for the anisotropy energy per Fe ion, or in terms of M_0, the magnetization of the sublattice at $0°$K, for the corresponding quantity per Fe ion

$$\kappa_1-\kappa_2=7.6(M/M_0)^2 \text{ cm}^{-1}, \qquad (4\cdot23)$$

which is in good agreement with the experimental value $5.3(M/M_0)^2$ cm^{-1}, given in § 2.

§ 5. Discussions

We will give several comments on the preceding calculation.

1) We have neglected the anisotropy energy due to the interaction between magnetic dipoles. It can be estimated from the work of Keffer[2] for MnF$_2$ as $0.2 (M/M_0)^2$ cm^{-1} per ion, which is smaller than $1/10$ of the value derived in the preceding section.

2) We have also neglected the anisotropy due to the anisotropic exchange interaction. This term would appear in the third order perturbation and contribute to the anisotropy energy in the order of $J\gamma_i^2$, the magnitude of which is much smaller than he quantity derived from the second order perturbation of the spin-orbit interaction. This fact suggests that the convergency of our perturbation calculation would be fairly well.

3) We have assumed that all Fe ions are in the ground state and thermal excitation into the higher states was neglected. This assumption would hold for the states of ψ_1, ψ_2 and ψ_h, while it may be uncertain for the state ψ_a, which is not so high ($E_a-E_0=-2.4\lambda=k\times340$ $°$K). But, as the experiment by Stout and Matarrese was operated at temperatures below $300°$K, our assumption would not involve fatal deficiency. The discrepancy between the experimental value $5.3(M/M_0)^2$ cm^{-1} and the theoretical one $7.6(M/M_0)^2$ cm^{-1} might result mainly from the above crude assumption.

4) We have treated the FeF$_2$ as a perfectly ionic crystal, and further neglected the states other than ^5D. Inclusion of these states and states arising from covalent FeF binding, of which both would have smaller spin multiplicity than 5, will not much influence our conclusion, but only serve to introduce superfluous parameters.

5) We have calculated the anisotropy energy under the assumption that the crystal-

line electric field potential on the ion exerted by the surrounding ions can be written in the quadratic form such as (3.1), and the coefficients A and B are determined from the comparison of the experimental g value with the theoretical value. Using (3.2a), (4.13) and (4.14), this potential for the $[a]$ site is given by

$$V = (36.2\xi^2 - 40.8\eta^2 + 4.6\zeta^2) \frac{2 \times 10^{-14}}{\langle r^2 \rangle} \ \text{erg.} \qquad (5 \cdot 1)$$

Since the Goldschmidt's radius of Fe^{++} is $0.83 A^{10)}$, $\langle r^2 \rangle$ may be in the order of $0.7 \times 10^{-16} \ cm^2$. Then, (5.1) can be written as

$$V = 3 \times 10^2 (36.2\xi^2 - 40.8\eta^2 + 4.6\zeta^2) \ \text{erg.} \qquad (5 \cdot 2)$$

Perhaps the first coefficient in the right hand side may be greater than 3×10^2.

On the other hand, we can estimate theoretically the cryatalline electric field potential under the assumption of point charges for both of Fe and F ions. According to the method of Evjen the calculation of the second order terms in the crystalline potential gives

$$V = 2.3 \times 10^3 \ (40\xi^2 - 12\eta^2 + 4\zeta^2) \ \text{erg,} \qquad (5 \cdot 3)$$

in the first approximation. The magnitude of (5.3) is about ten times larger than that of (5.2). This may be resulted from by the rough estimation of (5.3), i.e. first, in the first approximation of the method of Evjen, we have considered only the unit cell surrounding the Fe ion. But if we consider much more cells surrounding the Fe ion, the potential may be smaller than (5.3), because the influence of the remote ions would make the second order terms of the potential smaller. Secondly, we have neglected the induced polarization in (5.3), which cannot be neglected in the accurate calculation.

6) It is interesting that the energy levels splitted by the crystalline Stark effect are determined by the experiment of the g-factor and the magnetic anisotropy. It is desirable to confirm these levels by another measurement, say infrared absorption. In passing, we note that every transitions between those levels of FeF_2 correspond to the quadrupole radiation and are forbidden.

7) As the major part of the anisotropy energy comes from the second order perturbation of the spin-orbit interaction $\lambda(\boldsymbol{LS})$, the orientation of the spin $\boldsymbol{S}$ is favored parallel to the orbital angular momentum $\boldsymbol{L}$ than perpendicular to $\boldsymbol{L}$. On the other hand, the ground state is represented by the wave function $\psi_g = c_0\phi_0 + c_2\phi_2^+$, where $(c_0/c_2)^2 = 0.8$, i.e. in the ground state the orbital angular momentum $\boldsymbol{L}$ has a greater probability to point along ζ axis than along ξ or η axis. This is the physical interpretation of the origin of the anisotropy in FeF_2.

8) Finally, we will note that the anisotropy energy in FeF_2 is much larger than that in other substances. For ferromagnetic metals Fe and Ni the quantity is expressed in the form of $K_1(\alpha^4 + \beta^4 + \gamma^4) + K_2\alpha^2\beta^2\gamma^2$ (α, β and γ are direction cosines), and for C_0 K_1'. $\sin^2\theta + K_2'' \sin^4\theta$, where θ is the angle between the magnetization and the crystal axis. As the crystal structures of Fe and Ni are cubic, the second order terms such as K_1 and K_2 in FeF_2 do not exist, and further, their K_1 and K_2 are temperature dependent. Therefore,

it is difficult to make a comparison with FeF_2, but we give the value of Λ_1'' in Table II[11]. The anisotropy energy in antiferromagnetic substance $CuCl_2 \cdot 2H_2O$ is expressed as

Table II. Coefficient of the anisotropy energy at room temperature.

Fe	$\Lambda_1' = 4 \times 10^5$ erg/cc	$\kappa_1 = 2.5 \times 10^{-2}$ cm^{-1}/atom
Ni	$\Lambda_1' = -5 \times 10^4$	$\kappa_1 = 0.28 \times 10^{-2}$
Co	$\Lambda_1'' = 4 \times 10^6$	$\kappa_1' = 22 \times 10^{-2}$

Table III. Coefficient of anisotropy energy at low temperature.

$CuCl_2 \cdot 2H_2O$	$\kappa_1 = 0.510 \times 10^{-2}$cm^{-1}/atom
	$\kappa_2 = 1.68 \times 10^{-2}$
MnF_2	$\kappa_1 - \kappa_2 = 0.2$

$\kappa_1 \beta^2 + \kappa_2$ per ion, and for MnF_2 it has the same form as for FeF_2: $\kappa_1(\alpha^2 + \beta^2) + \kappa_2 \gamma^2$. From Table II and III, it is seen that the anisotropy energy of FeF_2 which has $\kappa_1 - \kappa_2 = 5.3$cm^{-1}, is very large. The difference between the magnitude of the anisotropy in MnF_2, the origin of which is dipole-dipole interaction, and that in $CuCl_2 \cdot 2H_2O$ comes from the difference of the magnitude of the spin, i.e. for Mn^{++} spin is $5/2$, while for Cu^{++} it is $1/2$. And further, the smaller anisotropy energy in MnF_2 than in FeF_2 can be understood by the fact that the Mn ion is in ^5S state, where there is no spin-orbit interaction at all.

There remain interesting problems of the anisotropy energy of antiferromagnetic substances such as CoF_2, Fe-S systems and hematite and various ferrites.

References

1) J. W. Stout and L. M. Matarrese, Rev. Mod. Phys. **25** (1953), 338.
2) T. Moriya and K. Yosida, Prog. Theor. Phys. **9** (1953), 663.
3) F. Keffer, Phys. Rev. **88** (1953), 608.
4) R. A. Erickson, Phys. Rev. **90** (1953), 779.
5) J. H. Van Vleck, J. Chem. Phys. **9** (1941), 85
6) P. W. Anderson, Phys. Rev. **79** (1950), 705.
   J. S. Smart, Phys. Rev. **86** (1952), 968.
7) J. H. Van Vleck, *The Theory of Electric and Magnetic Susceptibilities*, Chapter 7 (Oxford Press, 1932).
8) W. Gordy, W. V. Smith and R. F. Trambarulo, *Microwave Spectroscopy* p. 225 (John Wiley and Sons, 1953).
9) M. H. L. Pryce, Phys. Rev. **80** (1950), 1107.
10) F. Seitz, *Modern Theory of Solids* p. 93 (Mc Graw-Hill, 1940).
11) R. Becker and W. Döring, *Ferromagnetismus* (Springer, 1939).

Conservation Laws in Classical and Quantum Physics*

Eugene P. Wigner

Palmer Physical Laboratory, Princeton University, Princeton, New Jersey

The relation between the conservation laws and the symmetry properties of space and time is examined. If the simplest possible equation of motion takes the place of Newton's equation, the usual conservation laws are not obtained. In the quantum physics, on the contrary, there is a superabundance of conservation laws and their physical significance is not obvious.

As will be outlined in the ensuing article by Murai, there are several reasons for investigating the consequences of an extension of the inhomogeneous Lorentz group to include the transition to coordinate systems in uniformly accelerated motion. The only reason for investigating the reduction of the inhomogeneous Lorentz group to be discussed below is one of curiosity and a desire to be reminded of the fundamental role which Newton's equation of motion continue to play in physics.

We shall consider the consequences of the simplest possible equation of motion: "Every body remains at rest if no force acts on it." The equations of motion which correspond to this modified law are

$$m_\alpha \dot{x}_\alpha = -\partial f/\partial x_\alpha; \quad m_\alpha \dot{y}_\alpha = -\partial f/\partial y_\alpha; \quad m_\alpha \dot{z}_\alpha = -\partial f/\partial z_\alpha. \tag{1}$$

This applies then to every particle α, the f is a kind of potential, the gradient of which determines, however, not the acceleration but the velocity. Naturally, the results which we may obtain from (1) have no physical significance. However, by exploring the conservation laws which (1) leads to, we shall see that, in mechanics, space symmety without the specific form of Newton's law does not lead to all conservation theorems which we could expect. In quantum theory, on the other hand, the equivalent of (1) will entail these conservation laws even though their interpretation will not be too obvious.

If (1) holds, two coordinate systems which are in uniform motion with respect to each other are not equivalent any more, and there is no way to introduce a theory of relativity if one starts with (1). However, the symmetry properties of space alone remain preserved, its homogeneity and isotropy, and so does the homogeneity of time. As a

---

result, we expect that there will be a conservation theorem for energy (due to the homo-
geneity of time) and a momentum and angular momentum law. These expectations are
fulfilled in our mechanics (1) only partially : if f is invariant with respect to displacements
a momentum law follows from (1) which states that the centre of mass of an isolated
system is at rest ;

$$m_1 x_1 + m_2 x_2 + \cdots + m_n x_n = \text{Const.} \tag{2}$$

Similar equations hold for the y and z components. However, there does not seem to be
an energy principle, nor a conservation law for angular momentum. This result does not
conflict with the results of Hamel and Engel[1]. The connection between conservation laws
and symmetry is based, in mechanics, on the Hamiltonian formulation, and the equations
of motion (1) do not allow this formalism.

The quantum equation which corresponds to (1) can be derived by an adaptation of
Ehrenfest's principle[2]. This demands that it shall follow from

$$\frac{\partial}{\partial t} \Psi (x_1, \cdots, z_n) = Q\Psi \tag{3}$$

that the motion of the centre of mass

$$\bar{x}_\alpha = \int |\Psi|^2 x_\alpha d\tau$$

($\int \cdots d\tau$ indicates integration over the whole configuration space) obeys the law (1) where
$\partial f / \partial x_\alpha$ is again replaced by its average. Hence

$$m_\alpha \frac{d\bar{x}_\alpha}{dt} = m_\alpha \int x_\alpha (\Psi^* Q\Psi + \Psi (Q\Psi)^*) d\tau = -\int \frac{\partial f}{\partial x_\alpha} \Psi^* \Psi d\tau. \tag{3a}$$

It follows from the requirement that the total probability $\int |\Psi|^2 d\tau$ be independent of
time, just as in ordinary quantum theory, that Q is skew hermitean. Hence, one can
transform the second member of (3a) into

$$m_\alpha \int \Psi^* [x_\alpha Q\Psi - Q(x_\alpha \Psi)] d\tau.$$

If this is to be equal to the right side of (3a) for all Ψ^*, one must have

$$x_\alpha Q\Psi - Q x_\alpha \Psi = -\frac{1}{m_\alpha} \frac{\partial f}{\partial x_\alpha} \Psi. \tag{3b}$$

This will be valid for all Ψ if the commutator

$$[Q, x_\alpha] = \frac{1}{m_\alpha} \frac{\partial f}{\partial x_d}. \tag{3c}$$

The most general skew hermitean Q which satisfies (3a) and the similar equations for
the y and z components is

$$Q = \sum \frac{1}{m_\alpha} \left(\frac{\partial f}{\partial x_\alpha} \frac{\partial}{\partial x_\alpha} + \frac{\partial f}{\partial y_\alpha} \frac{\partial}{\partial y_\alpha} + \frac{\partial f}{\partial z_\alpha} \frac{\partial}{\partial z_\alpha} + \frac{1}{2} \Delta_\alpha f \right) + ig \tag{4}$$

The real function g can be quite arbitrary except that it must be invariant with respect to displacements and rotations, i. e., must depend only on the distances of the particles.

One can easily formulate the quantum mechanical conservation laws on the basis of (3) and (4). The energy principle becomes

$$E = i\hbar \int \Psi^* \left[\sum \frac{1}{m_\alpha} \left(grad_\alpha f \cdot grad_\alpha \Psi + \frac{1}{2} \Psi \Delta_\alpha f \right) + ig \Psi \right] d\tau \; ; \tag{5}$$

the angular momentum

$$M_z = i\hbar \int \Psi^* \sum \left(x_\alpha \frac{\partial}{\partial y_\alpha} - y_\alpha \frac{\partial}{\partial x_\alpha} \right) \Psi \, d\tau. \tag{6}$$

This last expression is identical with the one of ordinary quantum mechanics.

However, the connection between conservation laws and symmetry is not unique in quantum theory either. In contrast to the situation in the non-quantum formulation, there is a superabundance of conservation laws. In addition to the quantum momentum

$$P_z = -i\hbar \int \Psi^* \sum \frac{\partial \Psi}{\partial x_\alpha} d\tau, \tag{7}$$

there is the analogue of (2)

$$\int (\sum m_\alpha x_\alpha) \Psi^* \Psi \, d\tau. \tag{7a}$$

The modified Newton's law (1), which was used above to demonstrate the directness of the connection, in quantum theory, between conservation laws and symmetry, is the simplest law of motion that one can imagine. As was mentioned already, this law of motion precludes introducing any principle similar to the principle of relativity. It may be worth while to remark that, in addition, the mechanical law of motion is in this case, strictly valid also in quantum theory. It follows from (3) and (4) also that

$$\frac{\partial |\Psi|^2}{\partial t} = \sum \frac{1}{m_\alpha} \left[\frac{\partial}{\partial x_\alpha} \left(\frac{\partial f}{\partial x_\alpha} |\Psi|^2 \right) + \frac{\partial}{\partial y_\alpha} \left(\frac{\partial f}{\partial y_\alpha} |\Psi|^2 \right) + \frac{\partial}{\partial z_\alpha} \left(\frac{\partial f}{\partial z_\alpha} |\Psi|^2 \right) \right] \tag{8}$$

which is the continuity equation for particles which have the velocity components $(\partial f / \partial x_\alpha) / m_\alpha$ when they are at the point described by the variables of f; (8) expresses the fact that the time rate of change of $|\Psi|^2$—which is the density of the particles—is equal to the negative divergence of the current. This last quantity is equal to the product of density, $|\Psi|^2$, and the velocity. It is well known that not only is the equation (8) a consequence of the equations of motion but that, conversely, the equations of motion (1) can be deduced from the validity of (8) for a set of particles with density distribution $|\Psi|^2$. Thus the substitution of the quantum equation (3) and (4) really did not invalidate the mechanical equation (1), a circumstance which we could have inferred from the absence of h in (4). The equation of motion (1) would have made the

physics of the past fifty years very much easier : they would have made it impossible to introduce the theory of relativity, and quantization of the equations would not have changed their physical content.   The only new feature which the quantum theory introduces is the complex phase of our wave function Ψ and it is questionable whether this quantity could be attributed any physical significance.   Since the quantum conservation laws (5), (6) and (7) are all based on this complex phase, and vanish for a real wave function Ψ, their physical interpretation is open to question.

The above example was meant as a warning against a facile identification of symmetry and conservation laws.   It reminds us that the Hamiltonian formulation is necessary for that connection to hold in ordinary mechanics and that while it is always possible in quantum theory to deduce conservation laws from a symmetry condition, the interpretation of these conservation laws, and their significance, might be quite problematical.

The fact that the quantization of the equations of motion would not lead to a real quantum theory might have been foreseen from the fact that the uncertainty principle is hardly compatible with (1).   If all the coordinates have sharp values, this holds also for the forces $-grad\ f$.   This is true also in current quantum theory.   However, in the present theory, the forces determine the velocities, rather than the accelerations, and these become determined also.

References

1) G. Hamel, ZS. f. Math. u. Phys. **50** (1940), 1.
   F. Engel, Nachr. Gesellsch. Wiss. Göttingen p. 270 (1916)
   E. Bessel Hagen, Math. Ann. **84** (1921), 258.
2) P. Ehrenfest, ZS. f. Phys. **45** (1927), 455.

On the Group of Transformations in Six-dimensional Space, II

— *Conformal Group in Physics* —

Yasuhisa MURAI

Saitama University, Urawa

(Received January 22, 1954)

The role of the conformal group in the theory of physics is emphasized for various reasons. The relation between this group and the group of transformations in six-dimensional pseudo-euclidean space is shown. And as an illustration the theory of the non-local field is viewed through a concept of conformal group and it is shown that though in this way one can obtain equations, which the internal wave function must satisfy, how to interpret them is yet open to question.

The group of transformations in the pseudo-euclidean space of six-dimensions, of which the author tried a classification of irreducible unitary representaions[1], has an intimate relation with the conformal group in the Minkowski space. Of the conformal group many words have been said by various authors in different ways. In addition to the works cited in Schouten's paper[2], recently many physicists took up the problem related to this group ; for example, by Bopp[3], Hill[4], Hoffman[5] and Ingraham[6]. We shall begin by stating why this group is considered here, from a stand point of the physicist who has in mind the aim to obtain a mass spectrum, an extent of a particle and what not.

§ 1. The role of the conformal group

(a) The relativistic wave equations are classified with the aid of the classification of irreducible representations of the inhomogeneous Lorentz group : this was done by Bargmann and Wigner[7]. The inhomogeneous Lorentz group has two invariants which may be interpreted as $(mass)^2$ and $(mass)^2 \cdot (spin \ angular \ momentum)^2$, but as Wigner[8] has shown it can not be assigned discrete values to $(mass)^2$, and only three classes $(0, \infty)$, 0 and $(0, -\infty)$ are distinguished. It is natural to suppose that the wave equations which we are aiming to obtain are also related to some group and if one wants further that the mass is not contained parametically in them but can be assigned to have a definite value, the inhomogeneous Lorentz group must be somewhat extended. The conformal group contains this as a subgroup.

(b) The conformal group is the most general extension of the inhomogeneous Lorentz group. This was shown by Segal[9] mathematically. With respect to this point we should like to call atention to what Hill[10] has obtained. He showed that this extension is unique if one requires that the extended group reduces to the ordinary classical limit for vanishingly

small velocities, i.e., that it reduces to the one which leaves the equation $d^3x/dt^3 = \text{invariant}$. He also showed that one can not continue the extension further in the same line, that is, the above extension is the most general one.*

(c)   The equation of motion of a charged particle is, if we take into account the radiation reaction force,[11]

$$m\ddot{z}_\mu = 2e^2/3 \cdot (\dddot{z}_\mu + \dot{z}^2 \dot{z}_\mu) + e\dot{z}_\nu F_\mu^{\ \nu}.$$

If we consider the problem in one spatial dimension, this is rewritten as follows,

$$m\frac{\ddot{z}_1}{\sqrt{1+\dot{z}_1^2}} = \frac{2e^2}{3} \frac{d}{d\tau} \frac{\ddot{z}_1}{\sqrt{1+\dot{z}_1^2}} + \frac{c\,\dot{z}_0 F_1^{\ 0}}{\sqrt{1+\dot{z}_1^2}},$$

or,

$$m\left(1 - \frac{2}{3}r_0 \frac{d}{d\tau}\right)\frac{\ddot{z}_1}{\sqrt{1+\dot{z}_1^2}} = cF_1^{\ 0}.$$

This equation may be intuitively interpreted as if the Lorentz force acting on a point of the world line of a particle gives also influence to the point which is apart from that point by an order of a classical electron radius (considering m to be a mass of electron) along the world line. In fact the relationship between the radiation reaction and the size of a particle has been treated by Bopp[13]. What we should like to stress is that the equation, $\ddot{z}_1/\sqrt{1+\dot{z}_1^2} = \text{Const.}$, expresses the hyperbolic motion (in the limit $c\to\infty$ the uniformly accelerated motion) and so the term $\frac{2}{3}r_0 \frac{d}{d\tau} \frac{\ddot{z}_1}{\sqrt{1+\dot{z}_1^2}}$ plays a role when a particle is deviated from uniformly accelerated motion and then this effect looks identical with respect to any reference system which is performing a uniformly accelerated motion with each other. In one word we may say that a particle is seen to be of a same shape by any observer on these reference systems. Therefore it seems natural to suppose that the equation of which the extension of a particle is deduced is invariant under the trans-formation which leads a reference system to another one moving with a constant accelera-tion. The minimum group which contains these three transformations and the inhomo-geneous Lorentz transformations is the conformal group. Precisely, in forming a group, in addition to these transformations two more transformations must be added ; the one is a time component of above three transformations and the other is a dilatation. It is a curious feature that the dilatation is introduced here. As for the physical meaning of this transformation we cannot say hastily : we can get rid of these two transformations in the non-relativistic limit $c\to\infty$.

§ 2.   The relation between two groups

Now we shall consider the relation between the rotations in the six-dimensional pseudo-

---

euclidean space and the conformal transformations in the Minkowski space. The fundamental forms in these spaces are

$$(\eta^1)^2 + (\eta^2)^2 + (\eta^3)^2 - (\eta^0)^2 + (\eta^5)^2 - (\eta^6)^2 \tag{1}$$

and

$$g_{\mu\nu}x^\mu x^\nu = (x^1)^2 + (x^2)^2 + (x^3)^2 - (x^0)^2, \tag{2}$$

respectively. In the Minkowski space the conformal transformations are composed of the following three basic transformations:

    i)   translations and rotations, (i.e., the inhomogeneous Lorentz transformations)

$$'x^\mu = u^\mu{}_\nu x^\nu + a^\mu, \tag{3}$$

    ii)   dilatation,

$$'x^\mu = cx^\mu, \tag{4}$$

    iii)   inversion,

$$'x^\mu = x^\mu / g_{\lambda\nu}x^\lambda x^\nu. \tag{5}$$

If we introduce η^μ, ξ and ζ related to x^μ as follows,

$$\eta^\mu / \xi = x^\mu, \quad \zeta / \xi = 1/2 \, g_{\mu\nu} \, x^\mu x^\nu, \tag{6}$$

(2) becomes $g_{\mu\nu}\eta^\mu\eta^\nu - 2\xi\zeta = 0$. Further introducing η^5 and η^6 in places of ξ and ζ,

$$\sqrt{2}\xi = \eta^6 + \eta^5, \quad \sqrt{2}\zeta = \eta^6 - \eta^5,$$

we have as an expression for the fundamental form,

$$g_{\mu\nu}\,\eta^\mu\eta^\nu + (\eta^5) - (\eta^6) = 0,$$

i.e., (1) = 0. In these variables the transformations (3), (4) and (5) are written as follows,

    i)

$$'\eta^\mu = u^\mu{}_\nu \, \eta^\nu + 1/\sqrt{2} \, a^\mu (\eta^5 + \eta^6),$$

$$'\eta^5 = \eta^5 - 1/\sqrt{2} \, g_{\mu\nu} u^\mu{}_\lambda a^\nu \eta^\lambda - 1/2 \sqrt{2} g_{\mu\nu} a^\mu a^\nu (\eta^5 + \eta^6),$$

$$'\eta^6 = \eta^6 + 1/\sqrt{2} g_{\mu\nu} u^\mu{}_\lambda a^\nu \eta^\lambda + 1/2 \sqrt{2} g_{\mu\nu} \, a^\mu a^\nu (\eta^5 + \eta^6),$$

    ii)

$$'\eta^\mu = \eta^\mu,$$

$$'\eta^5 = 1/2 \, (c + 1/c) \eta^5 - 1/2 (c - 1/c) \eta^6,$$

$$'\eta^6 = -1/2 (c - 1/c) \eta^5 + 1/2 (c + 1/c) \eta^6,$$

    iii) $'\eta^\mu = \eta^\mu, \quad '\eta^5 = -\eta^5, \quad '\eta^6 = \eta^6.$

If we consider the transformations to be infinitesimal:

$$u^\mu{}_\nu = \delta^\mu{}_\nu + \mathcal{E}^\mu{}_\nu, \quad a^\mu = \mathcal{E}^\mu, \quad c = 1 + \mathcal{E},$$

the transformations i) and ii) become

    i) $'\eta^\mu = \eta^\mu + \mathcal{E}^\mu{}_\nu \eta^\nu + \mathcal{E}^\mu{}_5 \eta^5 + \mathcal{E}^\mu{}_6 \eta^6,$ (7)

$$'\eta'^5 = \eta^5 - g_{\mu\nu}\,\varepsilon^\mu\eta^\nu, \quad '\eta^6 = \eta^6 + g_{\mu\nu}\varepsilon^\mu\eta^\nu,$$

ii) $\qquad\qquad '\eta'^\mu = \eta^\mu, \quad '\eta^5 = \eta^5 - \varepsilon\eta^6, \quad '\eta^6 = \eta^6 - \varepsilon\eta^5.$ $\qquad\qquad$ (8)

Therefore we see that the infinitesimal rotation operators in $(\eta^j - \eta^k)$- and $(\eta^i - \eta^0)$·plane (which are L_i and M_i in our previous paper) generate Lorentz transformation, that in $(\eta^5 - \eta^6)$·plane (N) the dilatation and the translations are generated by $p_\mu - N_\mu$, where p_μ and N_μ are the infinitesimal rotation operators in $(\eta^\mu - \eta^5)$- and $(\eta^\mu - \eta^6)$·plane respectively. In fact, for example, the infinitesimal translation in the i-th direction is

$$'\eta^i = \eta^i + \varepsilon\eta^5 + \varepsilon\eta^6, \quad '\eta^5 = \eta^5 - \varepsilon\eta^i, \quad '\eta^6 = \eta^6 + \varepsilon\eta^i, \quad '\eta^\mu = \eta^\mu\,(\mu \neq i),$$

and this is a sum of

$$\begin{cases} '\eta^i = \eta^i + \varepsilon\eta^5 \\ '\eta^5 = \eta^5 - \varepsilon\eta^i \end{cases} \text{and} \begin{cases} '\eta^i = \eta^i + \varepsilon\eta^6 \\ '\eta^6 = \eta^6 + \varepsilon\eta^i \end{cases},$$

the former being caused by $-\varepsilon N_i$ and the latter by εp_i. The fact that $p_\mu - N_\mu$ corresponds to $-i\partial/\partial x^\mu$ in the Minkowski space has been already shown by Dirac[14] through a straightforward calculation.

Now we have to consider $p_\mu + N_\mu$. (We could have exchanged the roles of ξ and ζ, then $p_\mu + N_\mu$ comes to generate translation, and this exchange corresponds to the inversion (8).) As an example we take $\varepsilon(p_1 + N_1)$. This operator transforms η to $'\eta$:

$$'\eta^1 = \eta^1 - \varepsilon\eta^5 + \varepsilon\eta^6, \quad '\eta^5 = \eta^5 + \varepsilon\eta^1, \quad '\eta^6 = \eta^6 + \varepsilon\eta^1, \quad '\eta^\mu = \eta^\mu \quad (\mu \neq 1).$$

It can be analyzed in the following way,

$$(\eta^1,\ \eta^5,\ \eta^6) \xrightarrow{\text{inversion}} (\eta^1,\ -\eta^5,\ \eta^6) \xrightarrow{\text{translation}} (\eta^1 - \varepsilon\eta^5 + \varepsilon\eta^6, -\eta^5 - \varepsilon\eta^1,\ \eta^6 + \varepsilon\eta^1)$$

$$\xrightarrow{\text{inversion}} (\eta^1 - \varepsilon\eta^5 + \varepsilon\eta^6,\ \eta^5 + \varepsilon\eta^1,\ \eta^6 + \varepsilon\eta^1).$$

Writing the series of transformations in x-coordinates, we get

$$'x^1 = \frac{x^1 + 1/2\varepsilon x_\lambda x^\lambda}{1 + \varepsilon x^1}, \quad 'x^\mu = -\frac{x^\mu}{1 + \varepsilon x^1} \quad (\mu \neq 1).$$

If we put $'x^i = 0$, then

$$x^1 + 1/2\varepsilon(x^{\bar 1} - x^{\bar 0}) = 0, \text{ i.e., } x' = 1/2\varepsilon x^{0^2} + \cdots,$$

that is, the origin of the primed system performs a hyperbolic motion. Thus $p_i + N_i$ generates transformation which leads a coodinate system to another one in a uniformly accelerated motion. We set

$$\pi_\mu = 1/\sqrt{2}(p_\mu - N_\mu), \quad \xi_\mu = 1/\sqrt{2}(p_\mu + N_\mu).$$

The commutation relations among L_s, M_s, π_μ, ξ_μ and N are

$$[M_{\mu\nu},\ M_{\rho\sigma}] = i(g_{\nu\rho}M_{\nu\sigma} - g_{\nu\rho}M_{\mu\sigma} + g_{\mu\sigma}M_{\rho\nu} - g_{\nu\sigma}M_{\rho\mu}),$$

$$[M_{\mu\nu},\ \pi_\rho] = i(g_{\mu\rho}\pi_\nu - g_{\nu\rho}\pi_\mu), \quad [\pi_\mu,\ \pi_\nu] = 0,$$

$$[M_{\mu\nu}, \, \xi_\rho] = i(g_{\mu\rho}\xi_\nu - g_{\nu\rho}\xi_\mu), \quad [\xi_\mu, \, \xi_\nu] = 0, \tag{9}$$

$$[\xi_\mu, \, \pi_\mu] = ig_{\mu\mu} N, \quad [\xi_\mu, \, \pi_\nu] = -iM_{\mu\nu},$$

$$[\pi_\mu, \, N] = i\pi_\mu, \quad [\xi_\mu, \, N] = -i\xi_\mu, \quad [M_{\mu\nu}, \, N] = 0,$$

where $M_{oi} = M_i$ and $M_{ij} = L_k$. There are three invariants which commute with these fifteen elements. They are

$$Q = L^2 - M^2 + 2(\overline{\pi_0 \xi_0} - \overline{\pi \cdot \xi}) - N^2, \tag{10}$$

$$R = -(L \cdot M)N + L \cdot (\pi\xi_0 - \xi\pi_0) + M \cdot (\overline{\xi \times \pi}), \tag{11}$$

$$W = -(L \cdot M)^2 - 2(L \cdot \pi)(L \cdot \xi) + 2(L\pi_0 - \overline{\pi \times M}) \cdot (L\xi_0 - \overline{\xi \times M})$$

$$+ (MN + \overline{\pi_0 \xi} - \overline{\xi_0 \pi})^2 - (LN - \overline{\xi \times \pi})^2, \tag{12}$$

where $\overline{AB}$ means $1/2(AB + BA)$.

§ 3. The application

As π_μ generates translation it may be natural to put

$$\pi_i = \frac{1}{i} \frac{\partial}{\partial x^i}, \quad \pi_0 = -\frac{1}{i} \frac{\partial}{\partial x^0}. \tag{13}$$

Then the commutation relations (9) allow us to assign to other operators the following differentiation operator forms,

$$N = ix^\mu \frac{\partial}{\partial x^\mu} + 2i, \quad L_i = \frac{1}{i}\left(x^j \frac{\partial}{\partial x^k} - x^k \frac{\partial}{\partial x^j}\right), \quad M_i = \frac{1}{i}\left(x^i \frac{\partial}{\partial x_0} + x^0 \frac{\partial}{\partial x^i}\right) \tag{14}$$

$$\xi_\mu = 1/2 \cdot x_\nu x^\nu \pi_\mu + x^\mu N.$$

These expressions are the same as what E. L. Hill has obtained for the infinitesimal operators of the group which leaves the quation of motion, the time rate of the change of acceleration $= 0$, invariant. We have merely added $2i$ to N to make it Hermitean. When these expressions are introduced into the forms of three invariants, Q, R and W become identically equal to -4, 0 and 0 respectively.

If one will go to the non-relativistic limit, among above operators π_0, M_i, ξ_i and ξ_0 should be replaced by π_0/c, cM_i, $c^2\xi_i$ and $c\xi_0$ respectively, before the passage to limit. And in the limit $c \to \infty$, we get operators generating the transformations under which the equation $d^3x/dt^3 = 0$ is invariant. Thus, as a result of an enlargement of the symmetry properties of space-time we have to give up Newton's equation of motion. Consequently we can not expect the usual correspondence between symmetry properties and conservation laws, as is shown by Wigner in the proceeding paper. In other words, we lose the reason to identify, for example, π_i with a physical quantity representing a linear momentum. The unclearness which is inherent in this situation is further augmented if we go over to the case of non-local field. We shall adopt the description of this field by Yukawa. Let $(x|U|x')$ denote the scalar field. Starting with the expressions (14), we assume

them to be operators applying to this function from left. Similarly we define operators applying to this function from right as follows,

$$(x|U|x')\pi_\mu = -\frac{1}{i}\frac{\partial}{\partial x_\mu'}(x|U|x'),$$

$$(x|U|x')L_i = -\frac{1}{i}\left(x'^j\frac{\partial}{\partial x'^k} - x'^k\frac{\partial}{\partial x'^j}\right)(x|U|x'),$$

$$(x|U|x')M_i = -\frac{1}{i}\left(x'^i\frac{\partial}{\partial x'^0} + x'^0\frac{\partial}{\partial x'^i}\right)(x|U|x'), \qquad (15)$$

$$(x|U|x')\tilde{\xi}_\mu = -\frac{1}{i}\left(1/2(x'^2 - x'^{02})\frac{\partial}{\partial x'_\mu} - x'^\mu x'^\nu\frac{\partial}{\partial x'^\nu} - 2x'^\mu\right)(x|U|x'),$$

$$(x|U|x')N = -\left(x'^\mu\frac{\partial}{\partial x'^\mu} + 2\right)(x|U|x').$$

They are obtained from (14) replacing x^μ by x'^μ and changing the order of operation and the sign before differential operator. Further we introduce external and internal coodinates X^μ and r^μ;

$$X^\mu = 1/2(x^\mu + x'^\mu), \qquad r^\mu = x^\mu - x'^\mu,$$

and write $U(X, r)$ for $(x|U|x')$. Then if we form $[\pi_i, U]$, for example, it becomes

$$[\pi_i, U] = \frac{1}{i}\frac{\partial}{\partial X^i}U(X, r).$$

Doing with all the operators in the same way, we get the following operator forms,

$$\pi_\mu = \frac{1}{i}\frac{\partial}{\partial X_\mu}, \qquad N = i\left(X^\mu\frac{\partial}{\partial X^\mu} + r^\mu\frac{\partial}{\partial r^\mu} + 4\right),$$

$$L_i = L_i^0 + l_i = \frac{1}{i}\left(X^j\frac{\partial}{\partial X^k} - X^k\frac{\partial}{\partial X^j}\right) + \frac{1}{i}\left(r^j\frac{\partial}{\partial r^k} - r^k\frac{\partial}{\partial r^j}\right),$$

$$M_i = M_i^0 + m_i = \frac{1}{i}\left(X^i\frac{\partial}{\partial X^0} + X^0\frac{\partial}{\partial X^i}\right) + \frac{1}{i}\left(r^i\frac{\partial}{\partial r^0} + r^0\frac{\partial}{\partial r^i}\right), \qquad (15)$$

$$\tilde{\xi}_i = \frac{1}{2i}\left(X_\mu X^\mu + \frac{1}{4}r_\mu r^\mu\right)\frac{\partial}{\partial X^i} - \frac{1}{4i}r^i r^\mu\frac{\partial}{\partial X^\mu} + X^i N - X^0 m_i + l_j X^k - l_k X^j,$$

$$\xi_0 = -\frac{1}{2i}\left(X_\mu X^\mu + \frac{1}{4}r^\mu r_\mu\right)\frac{\partial}{\partial X^0} - \frac{1}{4i}r^0 r^\mu\frac{\partial}{\partial X^\mu} + X^0 N - m_i X^i.$$

They are operators applying to $U(X, r)$ and satisfy the commutation relations (9). The expressions of L_i and M_i are the same as what Yennie[15] has taken as an angular momentum operator in non-local field theory. Because we are treating scalar field, the β part of his expression does not appear. Of course we could add a x-independent part to L_i and M_i in (14), which satisfies by itself the same commutation relations as $M_{\mu\nu}$'s, thereby $\tilde{\xi}_\mu$ too

being modified correspondingly.

In the limit $r^\mu \to 0$, (16) tends to (14). As a consequence of introducing the internal coordinates, of which the physical meaning is quite obscure, the operators take a very complicated expressions: especially ξ_μ, which contains the external coordinates in an unfamiliar way. It may be worth while to note here that three invariants, being commutative with π_μ, don't contain the external coordinates except in the form of differentiation operators. Therefore, the solution of the equations, $[Q(\partial/\partial X^\mu, \partial/\partial r^\mu, r^\mu) - q]\Psi = 0$, etc., take a form of product $\Psi(X, r) = \Phi(x)\chi(r)$, and we get equations which the internal wave function $\chi(r)$ must satisfy. As an illustration we shall consider of an invariant Q. Substituting (15) into (10) we have

$$Q = 1/4\,(r^2 - r^{02})\,(\pi_0^2 - \pi^2) + 1/2 \cdot (r \cdot \pi - r^0\pi_0)^2 + 2(l^2 - m^2)$$

$$+ (r^2 - r^{02})\,(\partial^2/\partial r^2 - \partial^2/\partial r^{02}) + 2(r\,\partial/\partial r + r^0\,\partial/\partial r^0). \qquad (17)$$

The value of Q is determined from the classification of the irreducible representations of the group. The solution can be put in the form $\exp(ik_\mu X^\mu)\chi(r)$ and if we borrow the usual correspondence between translation and momentum, we have on the rest system (assuming $k_0^2 - k^2 = \kappa^2 \neq 0$),

$$\left[\frac{1}{4}\,(r^2 + r^{02})\kappa^2 + 2(l^2 - m^2) + (r^2 - r^{02})\left(\frac{\partial^2}{\partial r^2} - \frac{\partial^2}{\partial r^{02}}\right)\right.$$

$$\left. + 2\left(r\,\frac{\partial}{\partial r} + r^0\,\frac{\partial}{\partial r^0}\right) - Q\right]\chi(r) = 0.$$

When $l^2 = 0$, with a suitable change of variables the equation can be easily separated and the solution is found to be of the form $J_\mu(\kappa\lambda e^t)\,J_\nu(\kappa\lambda e^{-t})/\lambda^2\cosh t$, where $\lambda^2 = r^2 - r^{02}$, $r = \lambda \cosh t$ and μ and ν are related to Q in some way. So, if we think that the extent of a particle is represented by that of $\chi(r)$, for example, by the first zero of $\chi(r)$, the product of the rest mass and the radius of particle can gain a definite value but each cannot separately. This fact could be inferred from the mode of appearance of κ and r^μ in the above equation.

Thus as far as we follow the usual correspondence between the symmetry and the conservation laws, it seems that one can not have a satisfactory result on the basis of the conformal group. In fact, if one aims to introduce a universal length or in other words to obtain a spectrum of rest mass, it seems at first sight contradictory to start with a conformal group, which contains a dilatation operator. However, we shall remember the circumstance under which the principle of relativity was postulated together with that of the invariance of the light velocity in vacuo: this last fact would be thought to be incompatible with the first postulate. If one makes a success in attributing right meaning to every operators, some new feature will appear in the property of space-time.

The author would like to express his heartfelt thanks to Dr. Wigner for his kind encouragement.

References

1) Y. Murai, Prog. Theor. Phys. **9** (1953), 147.

2) J. A. Schouten, Rev. Mod. Phys. **21** (1949), 421.

3) F. Bopp, Lecture given at the International Conference on Theoretical Physics, Kyoto, Sept. 1953.

4) E. L. Hill, Phys. Rev. **84** (1951), 1165.

5) B. Hoffmann, Phys. Rev. **89** (1953), 49, 52.

6) R. L. Ingraham, Nuovo Cim. **9** (1952), 886.

7) V. Bargmann and E. P. Wigner, Proc. Nat. Acad. Sci. **34** (1948), 211.

8) E. P. Wigner, Ann. of Math. **40** (1939), 149.

9) I. E. Segal, Duke Math. Journ. **18** (1951), 221.

10) E. L. Hill, loc. cit.

11) P. A. M. Dirac, Proc. Roy. Soc. **A167** (1938), 148.

12) A. Pais, Prog. Theor. Phys. **10** (1953), 457.

13) F. Bopp, Zeits. f. Naturforsch. **3a** (1948), 564.

14) P. A. M. Dirac, Ann. of Math. **37** (1936), 429.

15) D. R. Yennie, Phys. Rev. **85** (1952), 877.

Note on the Slow Electrons in a Polar Crystal

Eijirô HAGA

Department of Physics, Faculty of Science, University of Niigata

(Received February 15, 1954)

A new perturbation technique is developed to investigate the properties of a slow electron in a polar crystal at absolute zero. The low-lying energy levels in a conduction band are calculated up to the order of K^2 by the second order perturbation calculation, where $\hbar K$ is the total momentum of the system, *i. e.* that of the polaron. For $K=0$ the energy agrees nearly with that obtained by the variational method. In the neighbourhood of $K=0$ the effective mass becomes $m\left(1-\dfrac{\alpha}{6}\right)^{-1}$, α being the coupling constant between an electron and the longitudinal optical modes of lattice vibrations. It is also shown that for $K=0$ the eigenfuction and the eigenvalue of the total Hamiltonian obtained by the variational procedure coincide with those in the zero-th approximation of the present perturbation method. The second order wave function is constructed which enables us to calculate the electrostatic potential and the charge density due to the ionic polarization around the electron.

§ 1.   Introduction

When an electron in a polar crystal is excited to a low-lying energy level in the conduction band, it interacts strongly with the longitudinal optical modes of the lattice vibrations and moves through the crystal accompanied by a cloud of phonons around it. Fröhlich, Pelzer and Zienau[1] have calculated the effective mass of this so-called polaron by the use of a variational method and have shown that it is nearly equal to the mass of the free electron. From this they have denied the assumption for the existence of self trapped electrons, which Landau[2] had advocated.

Recently Lee, Low, Pines[3] and Yokota[4] have dealt with the problem of polaron by the use of a variational technique more elaborate than that of FPZ. Seeing that the momentum operator of the polaron commutes with the Hamiltonian of the system and is, therefore, a constant of the motion, they have carried out the unitary transformation which makes the momentum of the polaron a " c " number and then determined the second unitary transformation so as to minimize the diagonal part of the transformed Hamiltonian. In this way they have obtained the low-lying energy levels and so the effective mass $m_v{}^*=m\left(1+\dfrac{\alpha}{6}\right)$ of the polaron. Gurari[5] have treated the same problem by a variational method which is essentially the same as that of FPZ and Yokota and have lead to the same results as those of the latters.

Since these authors all have used the variational method, the energy levels obtained by them should necessarily be higher than the true values, as it is concluded from the variational principle. Yokota[4] have estimated the accuracy of his procedure by comparing

the energy levels obtained by the variational method applied to one dimensional system consisting of an electron and a mode of lattice vibration with the exact solution obtained by Gross[3] for the same system. He has shown that the energy vs momentum curve by the variational method lies always higher than that of the exact solution, as it should be so, with only exception at $K=0$, where the both curves coincide, and that the discrepancy between the two curves increases with increasing K values.

Although the variational method gives a very good appoximation for the low-lying energy values of the polaron, we cannot assert that its effective mass calculated by the variational method also gives a good approximation, since the latter is determined by the curvature of the energy vs momentum curve. As the energy vs momentum curve calculated by the variational method bends upwards more strongly than that of the exact solution and the effective mass of the polaron is inversely proportional to its curvature, the variational method should give for the effective mass of the polaron a value smaller than the true one. And the matter of affairs, it is supposed, will be the same also for the present three dimensional case.

To find the effective mass of the polaron in the neighbourhood of $K=0$, it is necessary to calculate its energy levels exactly only in the neighbourhood of $K=0$. Hence in this paper we assume for the polaron momentum $\hbar K$ that the condition $\hbar^2 K^2/2m < \hbar\omega$ is always satisfied, where m is the effective mass of the electron determined by its interaction with the periodic lattice field and ω is the frequency of the longitudinal optical modes of the lattice oscillations, and treat the problem by the use of the perturbation method regarding the terms containing the polaron momentum as a perturbation. Hence, in the neighbourhood of $K=0$, the low-lying energy levels of the polaron obtained by us should be right up to the order K^2. The present method yields the energy vs momentum curve which lies always lower than that by the variational method with only exception at $K=0$ where the both curves coincide, and, therefore, gives the effective mass of the polaron larger than that by the variational method.

We make the same assumptions for the system consisting of a slow electron and lattice oscilllations as those which FPZ have made. Then neglecting the zero point energy of lattice oscillations we have for the Hamiltonian of the system

$$H = \sum_w a^*{}_w a_w \hbar\omega + \frac{gi}{V^{1/2}} \sum_w \frac{1}{w}(a^*{}_w e^{-iw\cdot r} - a_w e^{iw\cdot r}) + \frac{p^2}{2m}, \qquad (1)$$

where r and p are the position and momentum operators of the electron, and

$$g = \frac{\hbar^2}{m}(\pi u^3 \alpha)^{1/2}, \quad \frac{\hbar^2 u^2}{2m} = \hbar\omega \text{ and } \alpha = \frac{me^2}{\hbar^2 u}\left(\frac{1}{n^2} - \frac{1}{\varepsilon}\right).$$

Here the dimensionless number α plays the role of a coupling constant for the electron-lattice interaction, n and ε are the optical index of refraction and the static dielectric constant, respectively, and a_w and $a^*{}_w$ are the annihilation and creation operators of the phonon with the wave vector w, satisfying the commutation relation

$$[a_w, a^*{}_{w'}] = \delta_{ww'}. \qquad (2)$$

It is easily seen that the total momentum operator of the system $p + \sum_w a^*{}_w a_w \hbar w$ commutes with the Hamiltonian H (eq. (1)) and hence is a constant of the motion. Carrying out the unitary transformation with

$$U_1(r) = \exp\left[i\left(K - \sum_w a^*{}_w a_w w\right) \cdot r\right] \tag{3}$$

so that the total momentum becomes a " c " number, the Hamiltonian H is transformed into H^* :

$$H^* = U_1^{-1} H U_1 = \sum_w a^*{}_w a_w \hbar w + \frac{gi}{V^{1/2}} \sum_w \frac{1}{w}(a^*{}_w - a_w)$$

$$+ \frac{\hbar^2}{2m}\left(K - \sum_w a^*{}_w a_w w\right)^2, \tag{4}$$

where $\hbar K$ is the momentum of the total system, *i.e.*, that of the polaron and now is a " c " number. In § 2 we divide the Hamiltonian H^* into unperturbed and perturbing parts and determine the effective mass of the polaron by finding the low-lying energy levels of the polaron up to the order of K^2 by the use of the second order perturbation calculation. This procedure is justifiable in the case where $\hbar^2 K^2 / 2m \cdot \hbar \omega$. In § 3 we construct the wave function of the polaron in the second approximation and then calculate the electrostatic potential and the charge density due to the ionic polarization induced by the electron.

§ 2. Energy levels

To apply the perturbation technique to the present problem, it will be convenient to divide the Hamiltonian, eq. (4), into the three parts :

$$H^* = H_1^* + H_2^* + H_3^*, \tag{5}$$

where

$$H_1^* = \frac{\hbar^2}{2m} \sum_w a^*{}_w a_w (w^2 + u^2) + \frac{gi}{V^{1/2}} \sum_w \frac{1}{w}(a^*{}_w - a_w) + \frac{\hbar^2 K^2}{2m}, \tag{6}$$

$$H_2^* = \frac{\hbar^2}{2m}\left\{\left(\sum_w a^*{}_w a_w w\right)^2 - \sum_w a^*{}_w a_w w^2\right\}, \tag{7}$$

$$H_3^* = -\frac{\hbar^2}{m} \sum_w a^*{}_w a_w K \cdot w . \tag{8}$$

Since we are interested only in the case where $K \ll u$, we see the following relations to take place :

$$K \cdot w \leq K^2 \ll u^2 \qquad \text{for } w \leq K ,$$

$$K \cdot w < w^2 \ll u^2 \qquad \text{for } u \gg w > K,$$

$$K \cdot w \ll w^2 \ll u^2 \qquad \text{for } u \gg w \gg K ,$$

$$\boldsymbol{K}\cdot w \leqslant w^2 \leqq n^2 \qquad \text{for } u \geqq w \gg K,$$

and

$$\boldsymbol{K}\cdot w \leqslant uw < w^2 \qquad \text{for } w > u. \tag{9}$$

Hence, a term with any w in $H_3{}^*$ comes out to be much smaller than the corresponding term in the first term of $H_1{}^*$, and we, therefore, can regard $H_3{}^*$ as a perturbing term for the unperturbed Hamiltonian $H_1{}^*$.

Now we can exactly solve the eigenvalue problem with the unperturbed Hamiltonian $H_1{}^*$, since $H_1{}^*$ is identical with the Hamiltonian for an assembly of harmonic oscillators whose origins of positional coordinates or momenta are displaced, plus an additional constant. To do this, it will be convenient to carry out the unitary transformation with the transformation function

$$U_2 = \exp\left[\sum_w (a^*{}_w f_w - a_w f^*{}_w)\right], \tag{10}$$

where

$$f_w = -\frac{2mgi}{V^{1/2}\hbar^2}\,\frac{1}{w(w^2+u^2)}. \tag{11}$$

Then the unperturbed Hamiltonian $H_1{}^*$ is transformed into

$$\mathscr{H}_1 = U_2{}^{-1}H_1{}^*U_2 = \frac{\hbar^2}{2m}\sum_w a^*{}_w a_w (w^2+u^2) - \frac{2mg^2}{V\hbar^2}\sum_w \frac{1}{w^2(w^2+u^2)} + \frac{\hbar^2 K^2}{2m}, \tag{12}$$

since

$$U_2{}^{-1}a_w U_2 = a_w + f_w \quad \text{and} \quad U_2{}^{-1}a^*{}_w U_2 = a^*{}_w + f^*{}_w. \tag{13}$$

$\mathscr{H}_1$ is identical with the Hamiltonian of an assembly of harmonic oscillators with the frequency $\hbar/2m\cdot(w^2+u^2)$ plus an additional constant. If we denote its eigenfunction and eigenvalue as $\mathit{\Psi}$ and E respectively, we have $H_1{}^*U_2\mathit{\Psi} = EU_2\mathit{\Psi}$ as the solution for $H_1{}^*$.

As seen later, the diagonal element of $H_3{}^*$ corresponding to the lowest energy state of $H_1{}^*$ vanishes. Hence, if the magnitude of the second order perturbation energy of $H_2{}^*$ is neglesibly small compared with the magnitude of the lowest eigenvalue of $H_1{}^*$, we may regard $H_2{}^*$ as the perturbing term for the unperturbed Hamiltonian $H_1{}^*$ as we are interested only in the lowest energy state of the polaron. In facts the correction term due to $H_2{}^*$ amounts only to 7 percent of the lowest eigenvalue of $H_1{}^*$ even when α is as large as 6 (see later). We are thus justified to treat $H_2{}^*$ as well as $H_3{}^*$ as a perturbing term for the unperturbed Hamiltonian $H_1{}^*$.

Now it is necessary to obtain the matrix elements of the perturbing term $H_2{}^*$ and $H_3{}^*$ with respect to the eigenfunctions of the unperturbed Hamiltonian $H_1{}^*$ in order to treat our problem by the perturbation method. In other words, it is necessary to obtain the matrix elements of the perturbing terms $H_2{}^*$ and $H_3{}^*$ transformed by the unitary transform U_2 with respect of the eigenfunctions of $\mathscr{H}_1$, since $(U_2\mathit{\Psi}_r, H_i{}^*U_2\mathit{\Psi}_s)$

$= (\Psi_r, U_2^{-1} H_i^* U_2 \Psi_s)$, $(i=2$ and $3)$, where Ψ_r and Ψ_s are any two eigenfunctions of $\mathcal{H}_1$.

We, therefore, seek at first the perturbing terms H_2^* and H_3^* transformed by the unitary transform U_2. We find

$$\mathcal{H}_2 = U_2^{-1} H_2^* U_2 = \frac{\hbar^2}{2m} \{\sum_w (a^*_w a_w + a_w f_w^* + a^*_w f_w) w\}^2$$

$$- \frac{\hbar^2}{2m} \{\sum_w (a_w^* a_w + a_w f_w^* + a^*_w f_w + f_w f_w^*) w^2\}, \qquad (14)$$

where the summation $\sum_w f_w f_w^* w$ in the first term is omitted, since f_w depends only on the magnitude of w and hence the summation $\sum_w f_w f_w^* w$ over all possible w unitary transform U_2. We vanishes, and

$$\mathcal{H}_3 = -\frac{\hbar^2}{m} \sum_w a^*_w a_w \boldsymbol{K} \cdot \boldsymbol{w} - \frac{\hbar^2}{m} \sum_w (a_w f_w^* + a^*_w f_w) \boldsymbol{K} \cdot \boldsymbol{w}. \qquad (15)$$

The second term of $\mathcal{H}_1$ can easily be calculated replacing the summation by the the integral :

$$\sum \{\cdots\} = \frac{V}{(2\pi)^3} \int_0^\infty \int_0^\pi \int_0^{2\pi} w^2 \sin\theta \, dw \, d\theta \, d\varphi \{\cdots\}, \qquad (16)$$

the upper limit of the integral with respect to w being able to be replaced by infinity without essential error. The integration is straightforward and we find

$$-\frac{2mg^2}{V\hbar^2} \sum_w \frac{1}{w^2(w^2+u^2)} = -\alpha\hbar\omega. \qquad (17)$$

When the polaron momentum vanishes, $\mathcal{H}_3$ which is linear with respect to $\boldsymbol{K}$ also vanishes. Now we denote the normalized eigenfunction of the unperturbed Hamiltonian $\mathcal{H}_1$ with no phonon present as ϕ_0. It is clear that ϕ_0 represents the eigenstate corresponding to the lowest eigenvalue, $-\alpha\hbar\omega$, of $\mathcal{H}_1$, that is to say, $-\alpha\hbar\omega$ and $U_2\phi_0$ are the eigenvalue and the corresponding eigenfunction of the Hamiltonian H_1^* for $K=0$. It should be noted that the lowest energy and the corresponding wave function at $K=0$ obtained by the variational method, which LLP and Yokota have employed, coincide just with the zero-th approximate solution $-\alpha\hbar\omega$ and $U_2\phi_0$ obtained by our perturbation. method. Let the normalized wave function in which n_w phonons with wave vector w are present and no others be denoted as $\phi(n_w)$ and that in which n_w phonons with wave vector w and $n_{w'}$ phonons with wave vector w' are present and no others as $\phi(n_w, n_{w'})$, and so on.

We are now going to calculate the perturbation energy due to the perturbing term $\mathcal{H}_2 + \mathcal{H}_3$. Operating $\mathcal{H}_2$ to ϕ_0, we have

$$\mathcal{H}_2 \phi_0 = \frac{\hbar^2}{2m} \sum'_{w, w'} f_w f_{w'} \boldsymbol{w} \cdot \boldsymbol{w'} \phi(1_w, 1_{w'}), \qquad (18)$$

where the summation is taken over all w and w' except $w = w'$. The diagonal element of $\mathcal{H}_2$ with respect to ϕ_0 i. e. $(\phi_0, \mathcal{H}_2 \phi_0)$ vanishes, as easily seen from eq. (18), and hence the first order perturbation energy due to $\mathcal{H}_2$ vanishes. In eq. (18) the term $\sqrt{2}\, \hbar^2/2m \cdot \sum_w f^2_w w^2 \phi(2_w)$ is omitted, since it is neglesibly small if the volume of the crystal is taken to be large enough. In the followings we will always omit such terms that beome neglesibly small with increasing crystal volume. Operating $\mathcal{H}_3$ to ϕ_0, we have

$$\mathcal{H}_3 \phi_0 = -\frac{\hbar^2}{m} \sum_w f_w \boldsymbol{K} \cdot \boldsymbol{w}\, \phi(1_w), \tag{19}$$

The diagonal element of $\mathcal{H}_3$ with repect to ϕ_0, i. e. $(\phi_0, \mathcal{H}_3 \phi_0)$ vanishes as easily seen from eq. (19) and hence the first order perturbation energy due to $\mathcal{H}_3$ vanishes.

Now we are ready to calculate the second order perturbation energy:

$$\varDelta E = -\sum'_n \frac{|(\mathcal{H}_2 + \mathcal{H}_3)_{0n}|^2}{E_n - E_0}. \tag{20}$$

As f_w is a pure imaginary number, $(\mathcal{H}_2)_{0n}$ and $(\mathcal{H}_3)_{0n}$ are respectively real and pure imaginary as seen from eqs. (18) and (19). Hence eq. (20) can be written as

$$\varDelta E = -\sum'_n \frac{|(\mathcal{H}_2)_{0n}|^2}{E_n - E_0} - \sum'_n \frac{|(\mathcal{H}_3)_{0n}|^2}{E_n - E_0}. \tag{21}$$

The first and the second terms in eq. (21) shall be denoted as $\varDelta E_1$ and $\varDelta E_2$ respectively. Taking into account that $\phi(1_w, 1_{w'})$ and $\phi(1_{w'}, 1_w)$ represent quite the same state, we have from eq. (18)

$$(\phi(1_w, 1_{w'}), \mathcal{H}_2 \phi_0) = \frac{\hbar^2}{m} f_w f_{w'} \boldsymbol{w} \cdot \boldsymbol{w}'. \tag{22}$$

Inserting this into the first term in eq. (21) we have

$$\varDelta E_1 = -\frac{\hbar^2}{m} \sum'_{w, w'} \frac{|f_w f_{w'} \boldsymbol{w} \cdot \boldsymbol{w}'|^2}{w^2 + w'^2 + 2u^2}. \tag{23}$$

Replacing the summation by the integral we find

$$\varDelta E_1 = -\frac{2}{3} \left(\frac{1}{8} - \frac{1}{3\pi}\right) \alpha^2 \hbar \omega \sim -0.0126 \alpha^2 \hbar \omega, \tag{24}$$

which is 0.0126α times the zero-th approximation energy $-\alpha\hbar\omega$, amounting only to 7.5 percent of the latter even when α is as large as 6.

From eq. (19) we have

$$(\phi(1_w), \mathcal{H}_3 \phi_0) = -\frac{\hbar^2}{m} f_w \boldsymbol{K} \cdot \boldsymbol{w}. \tag{25}$$

Inserting this into the second term in eq. (21) we find

$$\Delta E_2 = -2 \frac{\hbar^2}{m} \sum_w \frac{|f_w \mathbf{K} \cdot \mathbf{w}|^2}{w^2 + u^2} = -\frac{\alpha}{6} \frac{\hbar^2 K^2}{2m} \,. \tag{26}$$

We are thus led to the expression for the polaron energy which is right up to the order of K^2. We have from eqs. (12), (17), (24), and (26)

$$E = -\alpha \hbar \omega - \frac{2}{3} \Big(\frac{1}{8} - \frac{1}{3\pi} \Big) \alpha^2 \hbar \omega + \frac{\hbar^2 K^2}{2m^*} \,, \tag{27}$$

where

$$m^* = m \Big/ \Big(1 - \frac{\alpha}{6} \Big) \tag{28}$$

means the so-called effective mass of the polaron. The effective mass value derived here is approximately equal to that derived by the variational method, $m_v^* = m(1 + \frac{\alpha}{6})$, only when $\alpha \leqslant 6$; and the ratio m^*/m_v^* increases with increasing α-value up to $\alpha = 6$. At $\alpha = 6$ m^* becomes infinity while m_v^* is 2; and when $\alpha > 6$ m^* becomes negative, the energy decreasing with increasing K. In the latter case the dependence of the polaron energy on its wave number K seems to be very complicated and the details of the dependence will be made clear if the polaron energy for K-values comparable with u is studied. It is supposed that as K increases from zero the polaron energy decreases and will reach to a minimal value E_0 at a certain K-value, K_0.

These circumstances have already been cleared up by Gross[6] for one-dimensional case. When the electron-lattice coupling is strong enough, according to him, the polaron energy decreases with increasing K-value near $K = 0$, though the effective mass is always positive, and reaches to its minimal value at a certain K-value and then increases. These abnormal behaviours of polaron energy when the coupling is as large as $\alpha > 6$ seem to correspond to Fröhlich[7]-Bardeen's[8] interpretation of superconducting state of metal.

It should be remarked that the energy levels lie lower than those derived by the variational method. It should be also remarked that they formally agree to those derived by the ordinary perturbation calculation up to the order of K^2, which are inadequate for such a strong electron-lattice coupling as in the present case, though the physical meaning of $\mathbf{K}$ is different in these two method, $\hbar \mathbf{K}$ meaning not the momentum of the total system, *i.e.*, that of the polaron, but the momentum of the electron only in the ordinary perturbation calculation.

For a typical polar crystal, NaCl, $\alpha = 5.2$, hence $m^* = 7.5\,m$ from eq. (28), while $m_v^* = 1.9\,m$. We see that the effective mass value by the present method is 4 times as large as that by the variational method in this case.

§ 3. Wave function

It will be also worth while to calculated the electrostatic potential and the charge density due to the ionic polarization around the electron in a polaron. For these purposes we must at first construct the wave function in the second approximation.

The wave function in the second approximation may be written in the form

$$\phi = \phi_0 + \phi_1 + \phi_2 , \tag{29}$$

where ϕ_0 is the unperturbed wave function, and ϕ_1 and ϕ_2 are the first and the second order corrections to it. According to the general scheme of the perturbation theory they are expressible as

$$\phi_1 = -\sum'_n \frac{(\mathcal{H}_2 + \mathcal{H}_3)_{n0}}{E_n - E_0} \phi_0 , \tag{30}$$

$$\phi_2 = \sum'_n c_n \phi_n ,$$

$$c_0 = -\frac{1}{2} \sum'_{n(\neq 0)} \frac{(\mathcal{H}_2 + \mathcal{H}_3)_{0n}(\mathcal{H}_2 + \mathcal{H}_3)_{n0}}{E_n^2} ,$$

$$c_n = \sum'_{k(\neq 0)} \frac{(\mathcal{H}_2 + \mathcal{H}_3)_{nk}(\mathcal{H}_2 + \mathcal{H}_3)_{k0}}{E_n \cdot E_k} \quad \text{for } n \neq 0 . \tag{31}$$

and

ϕ_1 may easily be expressed in terms of $\phi(1_w)$, $\phi(1_w, 1_{w'})$. We get using eqs. (18) and (19)

$$\phi_1 = 2 \sum_w \frac{f_w \boldsymbol{K} \cdot \boldsymbol{w}}{w^2 + u^2} \phi(1_w) - \sum'_{w,w'} \frac{f_w f_{w'} \boldsymbol{w} \cdot \boldsymbol{w}'}{w^2 + w'^2 + 2u^2} \phi(1_w, 1_{w'}). \tag{32}$$

To get the explicit expression for ϕ_2 we require all the non-vanishing matrix elements of $\mathcal{H}_2 + \mathcal{H}_3$ in the expressions for c_0 and c_n. To calculate them we need only to have $\phi(1_w, 1_{w'})$ and $\phi(1_w)$ operated by $\mathcal{H}_2$ and $\mathcal{H}_3$, as seen from eq. (18) and (19). The results are as follows:

$$\mathcal{H}_2 \phi(1_w, 1_{w'}) = \frac{\hbar^2}{2m} [2f^*_w f^*_{w'} \boldsymbol{w} \cdot \boldsymbol{w}' \phi_0 + 2 f^*_{w} \boldsymbol{w} \cdot \boldsymbol{w}' \phi(1_w) + 2 f^*_w \boldsymbol{w} \cdot \boldsymbol{w}' \phi(1_{w'})$$

$$+ 2 \boldsymbol{w} \cdot \boldsymbol{w}' \phi(1_w, 1_{w'}) + 2 \sum_{w''(\neq w)} f^*_{w'} f^*_{w''} \boldsymbol{w}' \cdot \boldsymbol{w}'' \phi(1_w, 1_{w''})$$

$$+ 2 \sum'_{w''(\neq w')} f^*_w f^*_{w''} \boldsymbol{w} \cdot \boldsymbol{w}'' \phi(1_{w'}, 1_{w''})$$

$$+ 2 \sum''_{w''(\neq w,w')} f_{w''} \boldsymbol{w} \cdot \boldsymbol{w}'' \phi(1_w, 1_{w'}, 1_{w''})$$

$$+ 2 \sum''_{w''(\neq w,w')} f_{w''} \boldsymbol{w}' \cdot \boldsymbol{w}'' \phi(1_w, 1_{w'}, 1_{w''})$$

$$+ \sum'''_{w''(\neq w,w')} \sum'''_{w'''(\neq w,w',w'')} f_{w''} f_{w'''} \boldsymbol{w}'' \cdot \boldsymbol{w}''' \phi(1_w, 1_{w'}, 1_{w''}, 1_{w'''}) , \tag{33}$$

and

$$\mathcal{H}_3 \phi(1_w) = -\frac{\hbar^2}{m} f^*_w \boldsymbol{K} \cdot \boldsymbol{w} \phi_0 - \frac{\hbar^2}{m} \boldsymbol{K} \cdot \boldsymbol{w} \phi(1_w)$$

$$- \frac{\hbar^2}{m} \sum_{w'} f_{w'} \boldsymbol{K} \cdot \boldsymbol{w}' \phi(1_w, 1_{w'}). \tag{34}$$

From eqs. (31), (33) and (34) we have as the second order correction to the wave function the following rather complicated expression

$$\phi_2 = -\frac{1}{2}\left[4\sum_w \frac{|f_w \mathbf{K}\cdot \mathbf{w}|^2}{(w^2+u^2)^2} + 2\sum'_{w,w'} \frac{|f_w f_{w'}\mathbf{w}\cdot \mathbf{w}'|^2}{(w^2+w'^2+2u^2)^2}\right]\phi_0$$

$$+4\sum_w \frac{f_w(\mathbf{K}\cdot \mathbf{w})^2}{(w^2+u^2)^2}\phi(1_w) + 4\sum'_{w,w'}\frac{f_w f_{w'} f^*_{w'}(\mathbf{w}\cdot \mathbf{w}')^2}{(w^2+u^2)(w^2+w'^2+2u^2)}\phi(1_w)$$

$$+4\sum'_{w,w'}\frac{f_w f_{w'}(\mathbf{K}\cdot \mathbf{w})(\mathbf{K}\cdot \mathbf{w}')}{(w^2+u^2)(w^2+w'^2+2u^2)}\phi(1_w,1_{w'})$$

$$+2\sum'_{w,w'}\frac{f_w f_{w'}(\mathbf{w}\cdot \mathbf{w}')^2}{(w^2+w'^2+2u^2)^2}\phi(1_w,1_{w'}) \tag{35}$$

$$+4\sum'_{w,w'}\sum'_{w''(\neq w)}\frac{f_w f^*_{w'} f_{w''} f^*_{w''}(\mathbf{w}\cdot \mathbf{w}'')(\mathbf{w}'\cdot \mathbf{w}'')}{(w^2+w'^2+2u^2)(w^2+w''^2+2u^2)}\phi(1_w,1_{w'})$$

$$+4\sum'_{w,w'}\sum''_{w''(\neq w,w')}\frac{f_w f_{w'} f_{w''}(\mathbf{w}\cdot \mathbf{w}')(\mathbf{w}\cdot \mathbf{w}'')}{(w^2+w'^2+2u^2)(w^2+w'^2+w''^2+3u^2)}\phi(1_w,1_{w'},1_{w''})$$

$$+\sum'_{w,w'}\sum''_{w''(\neq w,w')}\sum'''_{w'''(\neq w,w',w'')}$$

$$\times \frac{f_w f_{w'} f_{w''} f_{w'''}(\mathbf{w}\cdot \mathbf{w}')(\mathbf{w}''\cdot \mathbf{w}''')}{(w^2+w'^2+2u^2)(w^2+w'^2+w'^2+w'''^2+4u^2)}\phi(1_w,1_{w'},1_{w''},1_{w'''}).$$

In the above expression the summations in the coefficient of ϕ_0, the summation with respect to w' in the coefficient of $\phi(1_w)$, and the summation with respect to w'' in the coefficient of $\phi(1_w, 1_{w'})$ are straightforward if the summations are replaced by the integrals (see eq. (16)). Substituting the expression for f_w, eq. (11), into eqs. (32) and (35), we have

$$\phi_1 = -4\left(\frac{\pi u^3 \alpha}{V}\right)^{1/2} i \sum_w \frac{\mathbf{K}\cdot \mathbf{w}}{w(w^2+u^2)^2}\phi(1_w)$$

$$+4\frac{\pi u^3 \alpha}{V}\sum'_{w,w'}\frac{\mathbf{w}\cdot \mathbf{w}'\ \phi(1_w,1_{w'})}{ww'(w^2+u^2)(w'^2+u^2)(w^2+w'^2+2u^2)}, \tag{36}$$

and

$$\phi_2 = -\left\{\frac{\alpha}{24}\frac{K^2}{u^2} + \frac{1}{3}\left(\frac{7}{64} - \frac{1}{3\pi}\right)\alpha^2\right\}\phi_0 - 8\left(\frac{\pi u^3 \alpha}{V}\right)^{1/2} i \sum_w \frac{(\mathbf{K}\cdot \mathbf{w})^2}{w(w^2+u^2)^3}\phi(1_w)$$

$$-\frac{8}{3\pi}(\pi u^2 \alpha)^{3/2} i\frac{1}{V^{1/2}}\sum_w\left\{\frac{w}{2u(w^2+u^2)^3} + \frac{uw}{(w^2+u^2)^4} - \frac{w\sqrt{w^2+2u^2}}{(w^2+u^2)^4}\right\}\phi(1_w)$$

$$-16\frac{\pi u^3 \alpha}{V}\sum'_{w,w'}\frac{(\mathbf{K}\cdot \mathbf{w})(\mathbf{K}\cdot \mathbf{w}')}{ww'(w^2+u^2)^2(w'^2+u^2)(w^2+w'^2+2u^2)}\phi(1_w,1_{w'})$$

$$-8\frac{\pi u^3 \alpha}{V}\sum'_{w,w'}\frac{(\mathbf{w}\cdot \mathbf{w}')^2}{ww'(w^2+u^2)(w'^2+u^2)(w^2+w'^2+2u^2)^2}\phi(1_w,1_{w'})$$

$$+\frac{16}{3}\pi u^6 \alpha^2\frac{1}{V}\sum'_{w,w'}\frac{\mathbf{w}\cdot \mathbf{w}'}{ww'(w^2+u^2)(w'^2+u^2)(w^2+w'^2+2u^2)}\left\{\frac{1}{2u(w^2+u^2)}\right.$$

$$+ \frac{u}{(w^2+u^2)^2} - \frac{\sqrt{w^2+2u^2}}{(w^2+u^2)^2} \Bigg\} \phi(1_w, 1_{w'})$$

$$+ 32\left(\frac{\pi u^3 a}{V}\right)^{1/2} i \sum{}''_{w,w',w''}$$

$$\times \frac{(w\cdot w')(w\cdot w'')}{ww'w''(w^2+u^2)(w'^2+u^2)(w''^2+u^2)(w^2+w'^2+2u^2)(w^2+w'^2+w''^2+3u^2)} \cdot \frac{\phi(1_w, 1_{w'}, 1_{w''})}{}$$

$$+ 16\left(\frac{\pi u^3 a}{V}\right)^2 \sum{}'''_{w,w',w'',w'''}$$

$$\times \frac{(w\cdot w')(w''\cdot w''')}{ww'w''w'''(w^2+u^2)(w'^2+u^2)(w''^2+u^2)(w''^2+u^2)(w^2+w'^2+2u^2)}$$

$$\times \frac{1}{(w^2+w'^2+w''^2+w'''^2+4u^2)} \phi(1_w, 1_{w'}, 1_{w''}, 1_{w'''}), \tag{37}$$

where the second term in the coefficient of ϕ_0 is about $0.001a^2$ and is neglisibly small compared with unity even at $\alpha=6$.

Now we will ask for the expressions for the electrostatic potential and the ionic polarization charge density in a polaron. They may be calculated by the use of the second order perturbed wave function (eqs. (29), (36) and (37)). As the interaction term in our Hamiltonian (eq. (1)) is just the potential energy of the electron in the ionic polarization field, the coefficient of e in this term is the operator which represents the electrostatic potential at a point r. The mean value of this potential when the electron is at the point r_e is, therefore, expressible as

$$\varphi(r, r_e) = \frac{1}{e} \frac{gi}{V^{1/2}} \left((\psi(r_e), \left[\sum_w \frac{1}{w} (a^*_w e^{-iwr} - a_w e^{iw\cdot r}) \right] \psi(r_e) \right), \tag{38}$$

where the electronic wave function $\psi(r_e)$ is given by

$$\Psi(r_e) = U_1(r_e) U_2 \phi, \tag{39}$$

$U_1(r_e)$ being the unitary transform, eq. (3), with r replaced by r_e. Substituting eq. (39) into eq. (38), and taking into account the relations

$$U_1^{-1}(r_e) a_w U_1(r_e) = a_w e^{-iw\cdot r_e} \quad \text{and} \quad U_1^{-1}(r_e) a^*_w U_1(r_e) = a^*_w e^{iw\cdot r_e} \tag{40}$$

and eq. (13), we find

$$\varphi(r, r_e) = \varphi_0(r, r_e) + \varphi_1(r, r_e, K) \tag{41}$$

where

$$\varphi_0(r, r_e) = -\frac{gi}{V^{1/2}e} \left(\phi, \left[\sum_w \frac{1}{w} (f_w e^{iw\cdot(r-r_e)} - f^*_w e^{-iw\cdot(r-r_e)}) \right] \phi \right)$$

$$= -\frac{gi}{V^{1/2}e} \sum_w \frac{1}{w} (f_w e^{iw\cdot(r-r_e)} - f^*_w e^{-iw\cdot(r-r_e)}), \tag{42}$$

and

$$\varphi_1(\boldsymbol{r}, \boldsymbol{r}_e, \boldsymbol{K}) = -\frac{gi}{V^{1/2}e}\left(\phi, \left[\sum_w \frac{1}{w}\left(a_w e^{iw\cdot(r-r_e)} - a^*_w e^{-iw\cdot(r-r_e)}\right)\right]\phi\right). \quad (43)$$

Thus the ionic polarization charge density at $\boldsymbol{r}$ induced by an electron at $\boldsymbol{r}_e$ is

$$\rho(\boldsymbol{r}, \boldsymbol{r}_e) = -\frac{1}{4\pi}\nabla^2_r \varphi(\boldsymbol{r}, \boldsymbol{r}_e) = \rho_0(\boldsymbol{r}, \boldsymbol{r}_e) + \rho_1(\boldsymbol{r}, \boldsymbol{r}_e, \boldsymbol{K}), \quad (44)$$

where

$$\rho_0(\boldsymbol{r}, \boldsymbol{r}_e) = -\frac{gi}{4\pi V^{1/2}e}\sum_w w\left(f_w e^{iw\cdot(r-r_e)} - f^*_w e^{-iw\cdot(r-r_e)}\right) \quad (45)$$

and

$$\rho_1(\boldsymbol{r}, \boldsymbol{r}_e, \boldsymbol{K}) = -\frac{gi}{4\pi V^{1/2}e}\left(\phi, \sum_w w\left(a_w e^{iw\cdot(r-r_e)} - a^*_w e^{-iw\cdot(r-r_e)}\right)\right]\phi\right). \quad (46)$$

Substituting eq. (11) into eqs. (42) and (45) and relacing the summation by the integral (see eq. (16)), we have

$$\varphi_0(\boldsymbol{r}, \boldsymbol{r}_e) = -\frac{\alpha u \hbar^2}{me}\frac{1}{|\boldsymbol{r}-\boldsymbol{r}_e|}\left[1 - \exp(-u|\boldsymbol{r}-\boldsymbol{r}_e|)\right] \quad (47)$$

and

$$\rho_0(\boldsymbol{r}, \boldsymbol{r}_e) = -\frac{\alpha u^3 \hbar^2}{4\pi me}\frac{1}{|\boldsymbol{r}-\boldsymbol{r}_e|}\exp(-u|\boldsymbol{r}-\boldsymbol{r}_e|). \quad (48)$$

These are identical with the respective expressions of the electrostatic potential and the charge density for vanishing polaron momentum obtained by the variational method.

Substituting eqs. (29), (36) and (37) into eqs. (43) and (46), we have, together with eqs. (47) and (48), the electrostatic potential and the ionic charge density which are right up to the second order. It is early seen that $\varphi_1(\boldsymbol{r}, \boldsymbol{r}_e, \boldsymbol{K})$ and $\rho_1(\boldsymbol{r}, \boldsymbol{r}_e, \boldsymbol{K})$ are much smaller than $\varphi_0(\boldsymbol{r}, \boldsymbol{r}_e)$ and $\rho_0(\boldsymbol{r}, \boldsymbol{r}_e)$ respectively, and hence we need only preserve in ϕ, the second term of ϕ_2 (eq. (37)) together with ϕ_0:

$$\phi = \phi_0 - 8\left(\frac{\alpha u^3 \pi}{V}\right)^{1/2} i \sum_w \frac{(\boldsymbol{K}\cdot\boldsymbol{w})^2}{w(w^2+u^2)^3}\phi(1_w). \quad (49)$$

The other terms which have been neglected contribute nothing or much smaller values than the above two terms do.

Inserting eq. (49) into eqs. (43) and (46), and replacing the summation by the integral, we have

$$\varphi_1(\boldsymbol{r}, \boldsymbol{r}_e, \boldsymbol{K}) = -\frac{2\alpha u^3 \hbar^2}{\pi^2 me}\int dw\frac{(\boldsymbol{K}\cdot\boldsymbol{w})^2}{w^2(w^2+u^2)^3}\cos\left[w\cdot(r-r_e)\right] \quad (50)$$

and

$$\rho_1(\boldsymbol{r}, \boldsymbol{r}_e, \boldsymbol{K}) = -\frac{\alpha u^3 \hbar^2}{2\pi^3 me}\int dw\frac{(\boldsymbol{K}\cdot\boldsymbol{w})^2}{(w^2+u^2)^3}\cos\left[w\cdot(r-r_e)\right]. \quad (51)$$

The integrations are straightforward and we find

$$\varphi_1(\boldsymbol{r}, \boldsymbol{r}_e, \boldsymbol{K}) = \frac{12\alpha\hbar^2}{m e u^3}[\boldsymbol{K}\cdot(\boldsymbol{r}-\boldsymbol{r}_e)]^2 \left\{ \frac{1}{|\boldsymbol{r}-\boldsymbol{r}_e|^5} - \left(\frac{1}{|\boldsymbol{r}-\boldsymbol{r}_e|^5} + \frac{u}{|\boldsymbol{r}-\boldsymbol{r}_e|^4} \right.\right.$$

$$\left.\left. + \frac{u^2}{2|\boldsymbol{r}-\boldsymbol{r}_e|^3} + \frac{u^3}{6|\boldsymbol{r}-\boldsymbol{r}_e|^2} + \frac{u^4}{24|\boldsymbol{r}-\boldsymbol{r}_e|} \right) \exp\left(-u|\boldsymbol{r}-\boldsymbol{r}_e|\right) \right\}$$

$$-\frac{4\alpha\hbar^2 K^2}{m e u^3} \left\{ \frac{1}{|\boldsymbol{r}-\boldsymbol{r}_e|^3} - \left(\frac{1}{|\boldsymbol{r}-\boldsymbol{r}_e|^3} + \frac{u}{|\boldsymbol{r}-\boldsymbol{r}_e|^2} + \frac{u^2}{2|\boldsymbol{r}-\boldsymbol{r}_e|} + \frac{u^3}{8} \right) \exp\left(-u|\boldsymbol{r}-\boldsymbol{r}_e|\right) \right\},$$

$$(52)$$

which is finite at $\boldsymbol{r}=\boldsymbol{r}_e$, and

$$\rho_1(\boldsymbol{r}, \boldsymbol{r}_e, \boldsymbol{K}) = \frac{\alpha u^3 \hbar^2}{8\pi m e} \frac{[\boldsymbol{K}\cdot(\boldsymbol{r}-\boldsymbol{r}_e)]^2}{|\boldsymbol{r}-\boldsymbol{r}_e|} \exp\left(-u|\boldsymbol{r}-\boldsymbol{r}_e|\right)$$

$$-\frac{\alpha u^2 \hbar^2 w^2}{8\pi m e} \exp\left(-u|\boldsymbol{r}-\boldsymbol{r}_e|\right). \qquad (53)$$

eqs. (48) and (53) show that the induced charge density is flattened in the forward and the backward directions with respect to the motion of the polaron, as it has been shown also by the variatioal method.

References

1) H. Fröhlich, H. Pelzer, and S. Zienau, Phil. Mag. **41** (1950), 221.
2) L. Landau, Phys. Z. Sowjet. **3** (1933), 664.
3) T. D. Lee, F. E. Low and D. Pines, Phys. Rev. **90** (1953), 297.
4) T. Yokota, Busseiron Kenkyu, **64** (1953), 148.
5) M. Gurari, Phil. Mag. **44** (1953), 329.
6) E. P. Gross, Phys. Rev. **84** (1951), 818.
7) H. Fröhlich, Phys. Rev. **79** (1950), 845.
8) J. Bardeen, Phys. Rev. **80** (1950), 567.

The Generalization of Stueckelberg Formalism
in the Theory of Quantized Field*

Minoru HAMAGUCHI

Institute of Theoretical Physics, Kyoto University

Here is proposed a new attempt for overcoming the divergence difficulties in the interaction of elementary particles. For this purpose we introduce finite space time extension $\mathfrak{A}[x_\mu(\tau)]$ in each vertex of Feynman-Dyson-Graph, so that the usual S-matrix is transformed into the generalized form $S[\mathfrak{A}]$. $\mathfrak{A}$-extention is constructed by using the functionals $F[x_\mu(\tau)]$ with respect to proper time, and τ-region of its functional is taken analogously to that of Stueckelberg in his theory of finite time intervals. This method is first applied to the problem of self-energy of an electron, and the result is quite the same as introducing Feynman's cut-off procedure in the second order approximation, whereas in higher approximations our results differ definitely from that obtained by Feynman's cut-off method. Then, we propose a kinematical interpretation of the said finite extension of the region of interaction and further we investigate a possible generalization of the present method, which will finally be applied to the problem of anomalous magnetic moments of the nucleon.

§ 1. Introduction

There are two sorts of divergent results occuring in the quantum theory of elementary particles. These divergencies are : 1) the well known self-energy divergence[1], 2) the more serious boundary divergencies due to the sharp spatio-temporal limitation of the space-time extension $\mathfrak{A}[x_\mu]$ in which the collisions occur. It was pointed out by Stueckelberg[2),3)] that the later divergencies arise from region near the boundary, where processes occur without conservation of momentum-energy component normal to the super-surface. Moreover, he proposed that one will be able to obtain convergent results for this (so called) additional divergencies if the sharply defined boundary is transformed to the diffuse boundary. As is well known, it was the main object of his theory for finite time intervals to get convergent results for above divergencies which were different from the commonly encountered divergencies for infinite time intervals. The diffuse boundary is composed of the adjoining two supersurfaces. So, we can consider the finite space-time extension surrounded by these diffuse boundaries. Originally this extension is constructed so as to describe precisely the mechanism of the ideal experiment in the observations. It is also convenient that the characteristic behaviours of elementary particles are introduced by using this extension. On the other hand, if we consider this finite extension from a general idea of non-local theory it can be regarded as the spread of interacting elementary

---

* This was reported at the Kyoto meeting of the Physical Society of Japan, October, 31, 1952.

particles. As the consequence of the above reduction, we show here that one may remove not only additional divergencies but also the conventional ultra-violet catastrophe if we resort to the procedure which transforms the supersurface to the diffuse boundary or generally extends the points regions to the four dimensional finite regions.

Analogously to Stueckelberg's method, we can replace the δ-function in each vertex by the $\mathfrak{A}$-function without affecting the unitarity and causality of the array of probability amplitude forming the S-matrix`, but some generalized types introduced at the end of this paper may destroy the causality. $\mathfrak{A}$-function introduced above differs from that proposed by Stueckelberg in the theory of finite time intervals, because our function is expressed in a covariant form which has a finite space-time extension.

Since the equations of motion of elementary particles are represented in terms of the differential form[1] with respect to proper time, one may naturally be allowed to replace d_τ by $[\gamma_\mu \partial_\mu, \quad]$ (see (16)). In § 3, we know that this replacement plays a valuable role to interpret the velocity $d_\tau x_\mu$ from the quantum theoretical point of view.

The relativistic cut-off method introduced by Feynman[5] were the following : For the self energy problem, the $\delta(s_{12}^2)$ appearing in the action of interaction was replaced by a function of small width and great height. But its method was powerless for the problem of vacuum polarization, so he considered it analogously to the self energy problem and used a method of macking it convergent without spoiling the gauge invariancy. As a matter of fact, this method can be looked upon as the results of superposition of the effects of quanta of various mass. His method was, so to speak, merely a mathematical device that the above two problems were treated separately in this procedure, therefore we intend to construct a theory wherein these problems are treated from a unified point of view. But, the self energy calculation in our method introduce more complexity than Feynman's one, and the calculation of vacuum polarization is equivalent to Feynman's results in the second order approximation, but this equivalence does not hold for higher approximations.

In § 2, we used the fundamental idea that the probability amplitude is given by the functional derivative with respect to proper time $d_\tau F[x_\mu(\tau)]$, and that the finite space-time extension is constructed by the difference of these two functionals $F[x_\mu(\tau)]$. In the quantum theory, we think that the necessity of probability description might be inherent in the structure of matter. The description in terms of probability density may especially be needed for understanding phenomena associated with distance of the order 10^{-13} cm or less.

We lay stress on the fact that the $\mathfrak{A}[x_\mu(\tau)]$ exists only when the point lies inside a region (τ', τ''), and that it does not exist if τ lies outside it. Then we shall discuss the $\mathfrak{A}[x_\mu]$ and its more generalized spread. Since they are considered as the finite extension of elementary particles, we have the possibility of their interpretation from a kinematical standpoint in § 3. Consequently, the dynamical variables will be introduced as the variable in the ordinary δ-function in addition to the space time variables. The Fourier transforms of this spread play the role of convergence factors for the integral. Then, all the convergence factors will be unified into the type $\delta(x - x' - \lambda^n \cdot d^n x'/d\tau^n)$,

some of which are indispensable to the problem of anomalous magnetic moments of nucleon interacting with electromagnetic fields. Using the modified ∂-function which contains only the acceleration dependent term $\lambda^2 \cdot d^2x/d\tau^2$, it is evident that its interaction terms represents the anomalous magnetic moments of the nucleons according to the Foldy theory[3]. We concentrate our attention on this point and discuss the problem of nucleon charge in § 5. In § 4, we consider its relation to the theory of non-local action discussed by Pais and Uhlenbeck[12].

§ 2. Construction of the convergence factor

The rule in the quantum theory of elementary particles is given experimentaly by the transition probability such that the system is changed from a quantum distribution to an another quantum distribution.

On this occasion, we can consider the finite space-time region which means the existence of some effect in the macro-scopic matter fields according to the Stueckelberg's formalism. We also appeal to the spatio-temporal extension $\mathfrak{A}[x_\mu(\tau)]$ for describing this effect. Then the two distributions of quanta corresponding to the incident and the emergent wave packets will be considered in this extension.

For the convenience of the formulation, we introduce the following notations. The incident wave packets coming in the finite extension $\mathfrak{A}$ are described with φ which are composed of many wave packets $\varphi = \varphi', \varphi'', \varphi''', \cdots$. Then, an incident distribution indicating the number of quanta is simply written by $M(\varphi)$. Moreover, the amplitude characterizing the incident state is given by the functional $\Psi[M(\varphi)]$ which depends on the incident distribution $M(\varphi)$. The emergent state is given by the functional $\Psi[N(\chi)]$ of the emergent distribution $N(\chi)$ as well as the incident state. The quanta emerged from the finite space-time extension are also described with $\chi = \chi', \chi'', \chi''', \cdots$.

If we wish to interpret the S-matrix as an operator operating on a state vector Ψ, the state vector refers not to a time-like supersurface as in Schwinger theory because of the existence of the above extension $\mathfrak{A}[x_\mu]$, but to a diffuse boundary. Accordingly the transition amplitude must be given by the functional of the extension $\mathfrak{A}[x_\mu]$, the emergent distribution $N(\chi)$, and the incident distribution $M(\varphi)$,

$$S[\mathfrak{A}, N(\chi)/M(\varphi)] \rightarrow S[\mathfrak{A}] \qquad (1)$$

and transforms according to

$$\Psi'' = S[\mathfrak{A}]\Psi' \qquad (2)$$

the initial state into the final state. Considering the $\mathfrak{A}[x_\mu(\tau)]$ from a mathematical standpoint, it is a discontinuous function, with the two values zero or $\mathfrak{A}$ for the event point $x_\mu(\tau)$ outside or inside $\mathfrak{A}[x_\mu(\tau)]$. Furthermore, the extension $V[x_\mu(\tau)]$ is given by the difference between two functionals $F[x_\mu(\tau), \tau']$ and $F[x_\mu(\tau), \tau'']$.

$$V[x_\mu(\tau), \tau', \tau''] = F[x_\mu(\tau), \tau'] - F[x_\mu(\tau), \tau''] \qquad (3)$$

If we use the sign function $\varepsilon(\tau-\tau')$, the above functional $F[x_\mu(\tau),\tau']$ in detail is written as

$$F[x_\mu(\tau),\tau']=F[x_\mu(\tau)]\cdot\varepsilon(\tau-\tau'), \tag{3'}$$

and this equation have the two values $\pm 1/2\cdot F[x_\mu(\tau)]$ for $\tau\gtrless\tau'$. Then, $I[x_\mu(\tau),\tau',\tau'']$ in eq. (3) is a region surrounded by the two boundary layers which have the width $\Delta\tau'=\tau-\tau'$ and $\Delta\tau''=\tau-\tau''$, so it is written as $V[x_\mu(\tau)]$.

Now, all the physical quantities are written as the function of the variables $x_\mu(\tau)$ and τ according to the Feynman's five dimensional theory.[1] Therefore, it is very natural that the following two functionals can be introduced.

$$\left.\begin{array}{l} G[x_\mu(\tau),\tau]=F[x_\mu(\tau)]\cdot e^{-ik_5\tau} \\ \eta[x_\mu(\tau),\tau]=\delta[x_\mu(\tau)]\cdot e^{-ik_5\tau} \end{array}\right\} \tag{4}$$

We set up here the serious assumption that the later of the above equations is given by the functional derivative of the former, namely

$$\eta[x_\mu(\tau),\tau]=d_\tau G[x_\mu(\tau),\tau]. \tag{5}$$

This relation is one of the fundamental idea which is used in this paper. According to Appendix 1, $\eta[x_\mu(\tau),\tau]$ is reduced to the following form

$$\eta[x_\mu(\tau),\tau]\rightarrow(\gamma_\mu\partial_\mu+\partial_\tau)G[x_\mu(\tau),\tau], \tag{6}$$

where $\gamma_\mu\partial_\mu=\gamma_\mu\cdot\partial/\partial x_\mu$ and $\partial_\tau=\partial/\partial\tau$. If we drop the term $\exp(ik_5\tau)$ in eq. (6)*, we have

$$F[x_\mu(\tau)]=(2\pi)^{-4}\int(dk)\left(\frac{-i}{\gamma_\mu k_\mu+k_5}\right)\exp(ik_\mu\cdot x_\mu(\tau)). \tag{7}$$

Using the Fourier transform of the ε-functions in eq. (3), then we have

$$\varepsilon(\tau-\tau')-\varepsilon(\tau-\tau'')=(2\pi i)^{-1}\int_{-\infty}^{+\infty}\frac{d\alpha}{\alpha}\left[e^{i\alpha(\tau-\tau')}-e^{-i\alpha(\tau-\tau'')}\right]. \tag{8}$$

Therefore, the region $I[x_\mu(\tau)]$ in eq. (3) will exist when a point τ lies inside a region τ_0. If we fixed that $\tau_0=\tau'-\tau''$ was nearly equal to 10^{-13} cm, and that $\tau-\tau'=\Delta\tau'$, $\tau-\tau''=\Delta\tau''$, $\tau_0\gg\Delta\tau'$, $\Delta\tau''\gg\kappa^{-1}=\lambda$ (λ is a radius parameter of elementary particles.), the region $V[x_\mu(\tau)]$ becomes to the following form :

$$V[x_\mu(\tau)]=(2\pi)^{-4}\int(dk)\left(\frac{-i}{\gamma_\mu k_\mu+k_5}\right)\exp(ik_\mu\cdot x_\mu(\tau)) \tag{9}$$

$$, \text{ for } (\tau'<\tau<\tau'').$$

Putting $k_5=im$, $m=\sqrt{\kappa^2+\lambda^2}$, the Fourier amplitude of $I[x_\mu]$ then is given by

---

* This is similar to the procedure that we obtain the four dimensional eq. $(\gamma_\mu\partial_\mu-im)\psi(x)=0$ from the five dimensional wave eq. $\Gamma\psi(x,\tau)=\partial_\tau\psi(x,\tau)$, where $\Gamma\equiv\gamma_\mu\partial_\mu$, $\psi(x,\tau)=\psi(x)\cdot e^{-im\tau}$.

$$V_\lambda[k_\mu] = \left(\frac{-i}{\gamma_\mu k_\mu + i\sqrt{\kappa^2 + \lambda^2}} \right). \tag{10}$$

The $\mathfrak{A}[k_\mu]$ required to the formulation of S-matrix is given by

$$\mathfrak{A}[k_\mu] = \int_0^\infty V_\lambda[k_\mu] \cdot Q(\lambda) \, d\lambda, \tag{11}$$

with

$$Q(\lambda) = -\lambda G(\lambda). \tag{12}$$

$G(\lambda)$ is a any smooth function which have been discussed by Feynman[5] and Dyson[14] independently. In the conventional S-matrix elements in the momentum space representation, we replace the $\delta(k_\mu)$ in each vertex points by the above $\mathfrak{A}[k_\mu]$.

The mass-like term k_5 in eq. (4) has the arbitrary values, and we supposed that $F[x_\nu]$ was indirectly obtained from $\delta[x_\mu]$ in using the relation (5). In case of $k_5 = 0$, we assume that the same relation as (5) holds also between $\delta[x_\nu]$ and $F[x_\mu]$. In the above stand point, $\mathfrak{A}[x_\mu]$ corresponds physically to the covariant extension of elementary particles.

Then the n-th order contribution to the probability amplitude of a process according to Stueckelberg method[2] is the n-fold space time integral over $\mathfrak{A}$:

$$\varepsilon^n S_n[\mathfrak{A}](u'', \cdots \varphi'', \cdots / u', \cdots)$$
$$= i\varepsilon^n \int dx'' \mathfrak{A}[x''] \cdots \int dy'' \mathfrak{A}[y''] \cdots \int dx \mathfrak{A}[x] \cdots \int dx' \mathfrak{A}[x'] \cdots$$
$$\cdots u_A''^+(x'') \cdots \varphi_\alpha''^+(y'') \cdots \Delta_{B\ldots}^{(c)A\cdots\alpha\cdots}(x''-y'', \cdots, x''-x, \cdots, x''-x', \cdots)$$
$$\times u'^B(x') \cdots . \tag{13}$$

We represent the incoming waves of electrons, positrons, and photons by $u'(x)$, $v'(x)$ and $\varphi'(x)$, and the outgoing waves by their conjugate complex $u'''(x)$, $v'''(x)$ and $\varphi'''(x)$ in case of quantum electrodynamics. The causal function, $\Delta_{B\ldots}^{(c)A\cdots}(x''-y'', \cdots)$ is a covariant function of the n events. It is contragradient in its indices to the vector $\alpha\cdots\beta, \cdots$ or spinor $A\cdots B, \cdots$, indices of the packets.

§ 3. Kinematical interpretation*

It is a main object of the present investigation that the physical meanings of the region $\mathfrak{A}[x_\mu]$ and its generalized types, which are considered as the finite size of elementary particles, are given here by the dynamical treatment. In this article, we show that the special types of convergence factor in the non-local interactions are interpreted from the intuitional point of view. Now we interpret directly the convergence factor with the velocity or the acceleration, of the elementary particles.

The finite extension of elementary particles in the generalized Stueckelberg formalism in § 2, is briefly given by (see eq. (9))

---

* This was reported at the Tokyo meeting of the Physical Society of Japan, October, 16, 1953.

$$V[x_\mu(\tau)] = (2\pi)^{-4}(-i)P\int (dk)\frac{1}{\gamma_\mu k_\mu - im}\exp(ik_\mu \cdot x_\mu(\tau)). \tag{14}$$

According to the Schwinger's integral representation,[10] eq. (14) tend to

$$\left.\begin{aligned}V[x_\mu(\tau)] = \mathrm{const.}\int_{-\infty}^{+\infty}\frac{a}{|a|}c^{am}da\int (dk)c^{ik_\mu\cdot(x_\mu(\tau)+a\gamma_\mu)}, \\[2mm] \mathrm{const.} = -(2\pi)^{-4}/2.\end{aligned}\right\} \tag{15}$$

Now, the displacement from the space-time coordinates $x_\mu(\tau)$ to $x_\mu(\tau+\lambda)$ according to Feynman[4] is described as follows:

$$\left.\begin{aligned}\exp(\lambda H)x_\mu(\tau)\exp(-\lambda H) &= x_\mu(\tau+\lambda)\\[2mm] &= \sum_{n=0}^{\infty}\frac{\lambda^n}{n!}\frac{d^n x_\mu(\tau)}{d\tau^n}\\[2mm] &= x_\mu(\tau)+\lambda[H,\ x_\mu]+\frac{\lambda^2}{2!}[H,\ [H,\ x_\mu]]+\cdots,\\[2mm] H &= \gamma_\mu\partial_\mu \equiv \nabla.\end{aligned}\right\} \tag{16}$$

Then, an operator γ_μ in eq. (15) is expressed in the classical form[7][11]

$$[H,\ x_\mu]\equiv[\nabla,\ x_\mu]=\gamma_\mu\to dx_\mu/d\tau.$$

Accordingly, eq. (15) is equivalent to

$$V[x_\mu-x'_\mu(\tau)] = \mathrm{const.}\int_{-\infty}^{+\infty}\frac{a}{|a|}c^{am}da\cdot\delta\left(x_\mu-x'_\mu+a\frac{dx'_\mu}{d\tau}\right). \tag{17}$$

If we put $m=1/\lambda$, the Fourier amplitude in eq. (14) can be written over again as $\lambda/\lambda\gamma_\mu k_\mu - i$, so the above relation reduces to

$$V[x_\mu-x'_\mu]\sim\lambda\int_{-\infty}^{+\infty}\frac{a}{|a|}c^a da\cdot\delta\left(x_\mu-x'_\mu+a\lambda\cdot\frac{dx'_\mu}{d\tau}\right). \tag{18}$$

The a in (17) is a parameter with the dimension of length, on the contrary that in (18) has no dimensions. The procedure by which eq. (18) is obtained from eq. (14) is not the unique way, but it is interesting to note that the parameter a is independent of velocity dependent term $\lambda\cdot dx'/d\tau$.

In case that the Fourier amplitude is given by $1/\prod_{i=1}^{n}(\gamma_\mu k_\mu - im_i)$ instead of that in eq. (14), we use n-pieces of parameter $a_i(i=1, 2, \cdots\cdots n)$, and perform the n-folds integral with respect to a_i. Namely,

$$\int (dk)\cdot 1/\prod_{i=1}^{n}(\gamma_\mu k_\mu - im_i)\cdot\exp\{ik_\mu\cdot(x_\mu-x'_\mu)\}$$

$$\sim(-i/2)^n\int_{-\infty}^{+\infty}\frac{a_n}{|a_n|}da_n\cdots\int_{-\infty}^{+\infty}\frac{a_1}{|a_1|}da_1\cdot\exp\left(\sum_{i=1}^{n}a_i m_i\right)\cdot\delta\left(x_\mu-x'_\mu+\left[\sum_{i=1}^{n}a_i\right]\frac{dx'_\mu}{d\tau}\right)$$

$$\tag{19}$$

Also, in case of the Fourier amplitude $1/\prod_{i-1}^{n}(k^2+m_i^2)$, if $\prod_{i-1}^{n}(k^2+m_i^2)$ is linearized to $\prod_{i=1}^{n}(\gamma_\mu k_\mu+im_i)(\gamma_\mu k_\mu-im_i)$, we use $2n$-pieces of parameters a_i^+, a_i^-, $(i=1, 2, \cdots n)$. We therefore obtain the same expression as eq. (19).

We now take the following type omitting the parametric integral with respect to a,

$$
\begin{aligned}
\delta\left(x_\mu-x_\mu'+\lambda\frac{dx_\mu}{d\tau}\right)&=\delta(x_\mu-x_\mu'+\lambda\gamma_\mu)=(2\pi)^{-4}\int(dk)\exp\{ik_\mu\cdot(x_\mu-x_\mu'+\lambda\gamma_\mu)\} \\
&=(2\pi)^{-4}\int(dk)\exp\{ik_\mu\cdot(x_\mu-x_\mu')\}\cdot(1/e^{-i\lambda\tau_\mu k_\mu}), \\
1/e^{-i\lambda\tau_\mu k_\mu}&=1/1-(i\lambda\gamma_\mu k_\mu)+\cdots .
\end{aligned}
\tag{20}
$$

Comparing eq. (14) or (18) with eq. (20), we can interpret the parametric integrals with respect to a as a procedure which creates the finite series of eq. (14) from the infinite series of the denominator in eq. (20). In any case, the correlation between the cut-off of the exponential type in (20) and the cut-off in (14) will become clear if we carry out the parametric integrals $\int_{-\infty}^{+\infty}a/|a|\cdot f(a)da$ for eq. (20), where $f(a)$ is a function of a only.

Then, we can write all the cut-off factors as the modified δ-function in which the velocity dependent terms $dx_\mu/d\tau$ enter into the variables of it on the same footing with the space time coordinates x_μ, if possible the linearization of its factors.

Concerning the physical significance of $\delta(x_\mu-\lambda dx_\mu/d\tau)$, if it will be more generalized, let us permit to treat it as a type which contains not only the velocity $(dx_\mu/d\tau)$ but also the acceleration $(d^2x_\mu/d\tau^2)$ and etc.

Let us assume the action function for system of interacting spinor and electromagnetic fields in the following form:

$$
I_T=\int_R L(x')dx'+\int_S L(x)dx+\varepsilon\iint\bar{\psi}(x)\gamma_\mu\psi(x)V[x, x']A_\mu(x')dxdx'. \tag{21}
$$

$V[x, x']$ in the interaction term is a most generalized type which contains the $\lambda\cdot dx_\mu/d\tau$, $\lambda^2\cdot d^2x_\mu/d\tau^2$ and etc. Then we put ourselves in mind of the work discussed by Bohm, Weinstein and Kouts.[6][7] In this work, the current density $j_\mu'(x)$ is modified with the finite extension of moving particles.

The interaction term in (21) is the dynamical interaction which differs somewhat from the conventional interaction. Finally, it seems to us that such a dynamical interaction will be needed in the interaction of elementary particles.

§ 4. Field equations

The field equations derived from the variation of the action function (21) are

$$
\begin{aligned}
(\Box^2-\mu^2)A_\mu(x')&=\varepsilon\int\bar{\psi}(x)\gamma_\mu\psi(x)\delta(x-x'-\lambda\cdot dx'/d\tau)dx, \\
(\gamma_\mu\partial_\mu+\kappa)\psi(x)&=\varepsilon\gamma_\mu\int A_\mu(x')\delta(x-x'-\lambda\cdot dx'/d\tau)dx'.
\end{aligned}
\tag{22}
$$

$V'[x, x']$ in eq. (22) contains only the velocity dependent term as the present question. Considering eq. (22) as the modification of free field equations although in (22) the source of interaction was modified, then we have

$$e^{-\lambda(\gamma_\mu \partial_\mu)}(\Box^2-\mu^2)A_\mu(x')=\varepsilon\int \bar{\psi}(x)\gamma_\mu\psi(x)\partial(x-x')dx, \Bigg\}$$
$$e^{-\lambda(\gamma_\mu \partial_\mu)}(\gamma_\mu\partial_\mu+\kappa)\psi(x)=\varepsilon\gamma_\mu\int A_\mu(x')\partial(x-x')\psi(x)dx'. \Bigg\} \quad (23)$$

This eqs. (23) correspond to the infinite order equation discussed by Pais and Uhlenbeck. Accordingly, if we equalize eq. (22) to eq. (23), a source of the type $\partial(x-\lambda\cdot dx/d\tau)$ will be looked upon as an assembly which is composed of the infinite spinor fields.

In case of $V[x, x']$ which depends only on the acceleration i.e., $\partial(x-\lambda^2\cdot d^2x/d\tau^2)$, the field equations corresponding to (22), (23) are given respectively by

$$(\Box^2-\mu^2)A_\mu(x')=\delta(x-x'-\lambda^2\cdot d^2x'/d\tau^2), \quad (24)$$

and

$$e^{-\lambda(2/i)\sigma_{\mu\nu}\partial_\mu\partial_\nu'}(\Box^2-\mu^2)A_\mu(x')=\delta(x-x'), \quad (25)$$

with

$$\delta(x-x'-\lambda^2\cdot d^2x/d\tau^2)=\delta(x-x'-\lambda^2[\Gamma, [\Gamma, x]]) \quad (26)$$

$$=(2\pi)^{-4}\int (dk)\exp\{ik_\mu\cdot(x_\mu-x'_\mu-2i\lambda^2\sigma_{\nu\mu}\partial'_\nu)\},$$

$$[\Gamma, [\Gamma, x_\mu]]=[\gamma_\mu\partial_\mu, [\gamma_\mu\partial_\mu, x_\mu]]=2i\sigma_{\mu\nu}\partial_\nu.$$

Where, an operator $\sigma_{\mu\nu}$ is a Dirac's matrix spin tensor. It is evident that eq. (25) cannot indicate the non-local action which had been proposed by Pais and Uhlenbeck, on account of a matrix $\sigma_{\mu\nu}$.

Moreover, it is necessary to note that the exact correspondence between our theory and Pais-Uhlenbeck theory cannot be realized for using the V' which contains the third and higher derivatives of the space time coordinates x_μ with respect to proper time τ in addition to the space time coordinates. Finally, the complete coincidence between both theories is nothing but realized in the spinor fields.

§ 5.   Generalization of Delta functions

Let us consider the modified ∂-function which contains the infinite series consist of the high derivatives of coordinates x_μ with respect to proper time τ in addition to the space time coordinates.

$$\partial\left(x-x'-\sum_{n=1}^{\infty}\lambda^n\frac{d^nx'}{d\tau^n}\right)=\delta(x-x'(\tau+\lambda)) \quad (27)$$

In this equation, we used the following relation,

$$x'_\mu(\tau+\lambda)=\sum_{n=0}^{\infty}\lambda^n\frac{d^nx'_\mu}{d\tau^n}=x'_\mu+\sum_{n=1}^{\infty}\lambda^n\frac{d^nx'_\mu}{d\tau^n}.$$

Thus, the introduction of the cut-off factor with a form, $\sum_{n=1}^{\infty} \lambda^n \cdot d^n x / d\tau^n$, corresponds merely to the λ-displacement of proper time in the coordinates $x_\mu(\tau)$ which was discussed by Feynman, and the quantum mechanical description of its factor was already given by eq. (16).

But, we had hitherto (in § 3, § 4) discussed the cut-off factor containing one term which has the physical meaning of the velocity or the acceleration etc. That is to say, although it would be a simple procedure that the displacement of the coordinates $x_\mu(\tau)$ with respect to proper time was carried out, on the other side we see that the every term of the expansion of the coordinates $x_\mu(\tau+\lambda)$ with respect to λ have a dynamical significance in classical physics.

§ 6. Interactions with higher derivatives

According to the Foldy theory,[13] the frame work for describing the electromagnetic properties of Dirac particles is developed by finding the most general interaction terms consists of an infinite series involving arbitrary high derivatives of electromagnetic potentials evaluated at rhe position of the particle. The first term of this series represents the charge of the particle, the second its anomalous magnetic moments. Higher terms in the series describe the direct interactions of the particles.

It is the main object of our present investigation to prove a fact that the Foldy's interaction terms are contained within the non-local interaction terms introduced in § 5. Namely, the terms $dx/d\tau$, $d^2x/d\tau^2$, and $d^3x/d\tau^3$ etc. in § 5 correspond to the terms γ_μ, $(\gamma_\mu\gamma_\nu-\gamma_\nu\gamma_\mu)\partial_\nu$, and $(\gamma_\nu\partial_{o\mu}-\partial_{o\nu}\gamma_\mu)\partial_o\partial_\nu$, etc. respectively, owing to eq. (16).

On the other hand, the interaction terms investigated by Foldy have been generally expressed as a sum of the terms $\square^n\gamma_\mu l_\mu$ and $\square^n\gamma_\nu\gamma_\mu\left(\dfrac{\partial A_\mu}{\partial x_\nu}-\dfrac{\partial A_\nu}{\partial x_\mu}\right)$, where n is any non-negative integer.

Therefore, it will be evident that the Foldy's interaction is included in the interactions introduced in § 5. Naturally, our attentions turn into the second term $\gamma_\nu\gamma_\mu(\partial A_\mu/\partial x_\nu-\partial A_\nu/\partial x_\mu)$ of Foldy's interaction which describes the anomalous magnetic moments. This term is nothing else but the term discussed by Pauli[15]. For the interaction between Dirac particles and electromagnetic fields with the coupling $(\gamma_\nu\gamma_\nu-\gamma_\nu\gamma_\mu)\partial_\nu$ which is similar to the Pauli term, we have

$$\int \bar\psi(x)\gamma_\mu\psi(x)\,\delta(x-x'-\lambda^2\cdot d^2x'/d\tau^2)\,A_\mu(x')\,dx, \tag{28}$$

with

$$\delta(x-x'-\lambda^2\cdot d^2x'/d\tau^2)=(2\pi)^{-4}\int (dk)\exp\{ik_\mu\cdot(x_\mu-x_\mu'-\lambda^2(\gamma_\nu\gamma_\nu-\gamma_\mu\gamma_\nu)\partial_\nu')\}. \tag{29}$$

The end term $(\gamma_\nu\gamma_\mu-\gamma_\mu\gamma_\nu)\partial_\nu'$ in the exponential factor in (29) will operate to the electromagnetic potentials $A_\mu(x')$.

Putting the k_μ' for the energy momentum of photon which is transferred from the

Dirac particle, and performing the parametric integral with respect to a, eq. (29) becomes to the following form,

$$\int \frac{a}{|a|} da \cdot \delta(x - a\ddot{x}^2 \cdot d^2x'/d\tau^2) \sim \int (dk) \cdot 1/(2i) \lambda^2 \sigma_{\mu\nu} k_\mu k'_\nu \cdot \exp(ik \cdot x) \tag{30}$$

The term (30) is a cut-off factor of the new and eccentric type. Neversless the rationalization of this term is very easy (see Appendix II). Then, (30) is given by

$$\int (dk) \left[\frac{i\sigma_{\mu\nu} k_\mu k'_\nu}{2\lambda^2 \{(k \cdot k')^2 - k^2 k'^2\}} \right] \exp(ik \cdot x) \tag{31}$$

where a notation $(k \cdot k')$ is a scalar product of k and k'.

§ 7. Applications

The summaries of the application to quantum electrodynamics and to meson theories are given in this paragraph.

(A)   Application to quantum electrodynamics

According to the method introduced in § 2, we calculate the self energy of an electron and discuss the problem of vacuum polarization in second order approximation.

The $\mathfrak{A}[k_\mu]$ used in each vertex of Feynman-diagram is

$$\mathfrak{A}[k_\mu] = \int_0^\infty V_\lambda[k_\mu] \cdot Q(\lambda) \, d\lambda, \tag{32}$$

with

$$V_\lambda[k_\mu] = -i/\gamma_\mu k_\mu + i m_\lambda, \quad m_\lambda = \sqrt{\kappa^2 + \lambda^2} \quad \text{and} \quad Q(\lambda) = -\lambda \, G(\lambda).$$

In terms of $Q(\lambda)$ and $G(\lambda)$ the conditions are

$$\int_0^\infty Q(\lambda) \, d\lambda = 0, \qquad \int_0^\infty G(\lambda) \, d\lambda = 1. \tag{33}$$

Fig. 1.

electron   - - - - - - photon

Feynman diagram of the process contributing to the lowest order self-energy of an electron.

If we identify the two different values λ_1, λ_2 of $V_\lambda[k_\mu]$ which is used in stead of $\delta[k_\mu]$ in the two vertices, i.e. $\lambda_1 = \lambda_2 = \lambda$, the contributions of the two $V_\lambda[k_\mu]$ to the S-matrix element will arrange as a term $1/k^2 + \kappa^2 + \lambda^2$.

Then, the self energy integrals of an electron in the momentum space representation is given by

$$\frac{e^2}{\pi i} \int (dk) \frac{1}{(k^2 - \kappa^2 - \lambda^2)} \gamma_\alpha \frac{(k + \kappa)}{(k^2 - \kappa^2)} \gamma_\alpha \frac{1}{(p - k)^2}, \tag{34}$$

and a part of the integrand of its term can be written as

$$\frac{1}{(k^2 - \kappa^2)(k^2 - \kappa^2 - \lambda^2)} = \frac{1}{\lambda^2} \int_{\kappa^2}^{\kappa^2 + \lambda^2} (k^2 - L)^{-2} dL. \tag{35}$$

Thus, we can replace a part corresponding to the left hand side of the above equation by a term $1/\lambda^2 \cdot (k^2 - L)^{-2}$ and at the end integrate the results with respect to L from κ^2 to $\kappa^2 + \lambda^2$.

We can write the integral (34) as a form

$$\frac{e^2}{\pi i} \frac{1}{\lambda^2} \int_{\kappa^2}^{\kappa^2 + \lambda^2} dL \, \gamma_\alpha \int (dk)(k+\kappa)(k^2 - L)^{-2}(k^2 - 2pk + p^2)^{-1} \gamma_\alpha. \tag{36}$$

Owing to the Formula

$$a^{-2}b^{-1} = \int_0^1 dx \, (ax + b(1-x))^{-2},$$

this term is written over as follows:

$$\frac{e^2}{\pi i} \frac{1}{\lambda^2} \gamma_\alpha \int_0^1 2x dx \int_{\kappa^2}^{\kappa^2 + \alpha^2} dL \int (dk)(k+\kappa)(k^2 - 2pk(1-x) + p^2 - Lx + p^2x)^{-3} \gamma_\alpha. \tag{37}$$

Now, if the integrals with respect to k and L are furnished, we obtain the form

$$\frac{-e^2}{4\pi \lambda^2} \gamma_\alpha \int_0^1 dx \, (p - \kappa - px) \ln \left(\frac{p^2 x + \kappa^2 + \lambda^2 - p^2}{p^2 x + \kappa^2 - p^2} \right) \gamma_\alpha. \tag{38}$$

We integrate the rational function with respect to x from 0 to 1 and at the end execute the weighted integration with respect to λ, then we have

$$\frac{e^2}{8\pi} \left[(8\kappa - 2p) \ln \left(\frac{\kappa^2 + \lambda^2}{\kappa^2} \right) + 2 \left\{ (4\kappa - 2p) \left(\frac{\kappa^2 - p^2}{p^2} \right) - p \left(\frac{\kappa^2 - p^2}{p^2} \right)^2 \right\} \right.$$
$$\left. \times \left\{ \ln \left(\frac{\kappa^2 + \lambda^2}{\kappa^2 + \lambda^2 - p^2} \right) - \ln \left(\frac{\kappa^2}{\kappa^2 - p^2} \right) \right\} \right]. \tag{39}$$

When applied to a state of an electron of momentum p satisfying $pu = \kappa u$, it gives for the changes in mass

$$\Delta \kappa = \kappa \left(\frac{e^2}{4\pi} \right) \left\{ 3 \ln \left(\frac{\kappa^2 + \lambda^2}{\kappa^2} \right) \right\}. \tag{40}$$

In this case, the second term in the brackets in (39) does not contribute for the change in mass.

On the other hand, the results of the calculation of the vacuum polarization in second order approximation is entirely the same expression as the results obtained by Feynman.[5] Although we omit this calculations it is necessary to note that the results in higher approximations differ from that obtained by Feynman.

(B) Application to meson theories

The problem of the magnetic moments of the nucleon had been discussed by Case[9] several years ago. In this theory, to describe the meson field he chose the simplest possibility which may be expected to yield interesting results — namely the pseudoscalar theory. Of the meson-nucleon couplings for this case, the pseudoscalar coupling was chosen.
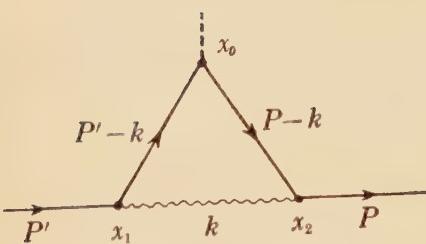
Fig. 2.

· nucleon ----- meson ~~~~ photon

Feynman diagram of the process contributing to the lowest order nucleon charge.

We also calculate the nucleon charge with the same description as Case's work, though the modified δ-functions are introduced in each vertex. Then if we replace a δ-function in a vertex point x_0 by a modified δ-function $\delta_c(x-\lambda^2\cdot d^2x/d\tau^2)$ and the two δ-functions in the vertex points x_1, x_2 by the same modified function $\delta_c(x-\lambda\cdot dx/d\tau)$, the meson-nucleon and nucleon-photon interacting Hamiltonians are given by

$$
\left.\begin{array}{l}
H_{NM}=if_v\bar\psi(x)\gamma_5\tau_v\psi(x)\,\delta_c(x-x'-\lambda\cdot dx'/d\tau)\phi(x'),\\[2mm]
H_N^{ert}=-ie\bar\psi(x)\dfrac{(1-\tau_3)}{2}\gamma_\mu\psi(x)\delta_c(x-x'-\lambda^2\cdot d^2x''/d\tau^2)A_\mu(x'),
\end{array}\right\} \quad (41)
$$

where the operators ψ, $\bar\psi$, ϕ, A_ν indicate the nucleons and its adjoints, the mesons and the photons. But the modified δ-functions $\delta_c(x-\lambda\cdot dx/d\tau)$ and $\delta_c(x-\lambda^2\cdot d^2x/d\tau^2)$ in (41) are

$$
\left.\begin{array}{l}
\delta_c(x-\lambda\cdot dx/d\tau)=\displaystyle\int\dfrac{a}{|a|}f(a)da\cdot\delta(x-a\lambda\cdot dx/d\tau),\\[3mm]
\delta_c(x-\lambda^2\cdot d^2x/d\tau^2)=\displaystyle\int\dfrac{a}{|a|}f(a)da\cdot\delta(x-a\lambda^2\cdot d^2x/d\tau^2).
\end{array}\right\} \quad (30')
$$

To the brevity of the actual calculations, in stead of the Fourier amplitudes of the modified δ_c-functions in eq. $(30')$, we use

$$
\int_0^\infty\frac{ia G(\alpha)}{\gamma_\mu k_\mu+i\sqrt{\kappa^2+\alpha^2}}da,
$$

and

$$
\frac{\gamma_\mu\gamma_\nu(P'-k)_\mu(P-P')_\nu}{\lambda^2(P'-k)^2(P-P')^2},
$$

respectively (see (11), (31)).

Now, the matrix element corresponding to Fig. II in the momentum space representation is given by

$$
H=\frac{-ieA_\mu}{8\hbar c\lambda^2}\left(\frac{-i}{2\pi}\right)\int(dk)\bar\psi(x_0)T
$$

$$
\times\;\frac{[i\gamma_\nu(P'-k)_\nu+\kappa_0]\gamma_\sigma(P'-k)_\sigma i\tau_\rho(\Delta P)_\rho\gamma_\mu[i\gamma_\lambda(P-k)_\lambda+\kappa_0]}{(k_\mu^2+\kappa^2+\alpha^2)[(k_\mu-P'_\mu)^2+\kappa_0^2](P'-k)^2(\Delta P)^2[(k_\mu-P'_\mu)^2+\kappa_0^2][k_\mu^2+\kappa^2]}\cdot\psi(x_0),
$$

$$
\tag{42}
$$

where $T=(f_3^2+f_4^2)((1-\tau_3)/2)+(f_1^2+f_2^2)((1+\tau_3)/2)$, $\Delta P=P-P'$. f_1 to f_4 are

the pseudoscalar coupling const. and τ are the conventional isotopic spin matrices that the eigenvalue $+1$ of τ_3 corresponds to neutrons, -1 to protons. κ_0 and κ are the nucleon and the meson masses, respectively. In the above equation, if we use the relation

$$\frac{1}{k^2+\kappa^2+\alpha^2}\cdot\frac{1}{k^2+\kappa^2}=-\frac{1}{\alpha^2}\int_{\kappa^2}^{\kappa^2+\alpha^2}(k^2+L)^{-2}dL,$$

it can be written as

$$H=\frac{-ieA_\mu}{8\hbar c\lambda^2}\left(\frac{i}{2\pi^4}\right)\left(\frac{1}{\alpha^2}\right)\int_{\kappa^2}^{\kappa^2+\alpha^2}dL\int(dk)\,\overline{\psi}(x_0)\cdot T$$

$$\times\frac{[i\gamma_\nu(P'-k)_\nu+\kappa_0]\,\gamma_3(P'-k)_\sigma\gamma_\rho\varDelta P_\rho\gamma_\mu[i\gamma_\lambda(P-k)_\lambda+k_0]}{(\varDelta P)^2(k^2+L)^2[(k-P')^2+\kappa_0^2](P'-k)^2[(k-P)^2+\kappa_0^2]}\psi(x_0). \tag{43}$$

In eq. (43), we have the use of a formula

$$a^{-2}b^{-1}c^{-1}d^{-1}=\int_0^1 dx\int_0^x dy\int_0^y dz\,(4z)\,[az+b(y-z)+c(x-y)+d(1-x)]^{-5}$$

and integrate the results with respect to k. Then if we select the terms which contribute to the anomalous magnetic moments of the nucleon, obtain finally the following results. (The detailed calculations are omitted.)

$$\frac{-ieA_\mu}{8\hbar c\lambda^2}\left(\frac{1}{12\alpha^2\pi^2}\right)\int_{\kappa^2}^{\kappa^2+\alpha^2}dL\int_0^1 dx\int_0^x dy\int_0^y dz\cdot(4z)\,\overline{\psi}(x_0)\,T$$

$$\times\frac{(z-x)(1-2z+zx)\kappa_0\sigma_{\mu\nu}\varDelta P_\nu}{\{(x-1)(x-z)(\varDelta P)^2-\kappa_0^2(z-1)^2+\kappa_0^2(x-y)-Lz\}^3}\cdot\psi(x_0) \tag{44}$$

At the end, we integrate (44) with respect to L from κ^2 to $\kappa^2+\lambda^2$. Then, this term is simplified to

$$\frac{-ieA_\mu}{48\pi^2\hbar c\lambda^2\alpha^2}\int_0^1 dx\int_0^x dy\int_0^y dz\,\overline{\psi}(x_0)\,T(z-x)(1-2z+zx)\kappa_0\sigma_{\mu\nu}\varDelta P_\nu$$

$$\times\left\{\frac{1}{[\kappa_0^2\{x-y-(z-1)^2\}-(\kappa_0^2+\alpha^2)z]^2}-\frac{1}{[\kappa_0^2\{x-y-(z-1)^2\}-\kappa_0^2z]^2}\right\}\psi(x_0). \tag{45}$$

Here the $(\varDelta P)^2$ in the denominator in (44) has been set equal to zero since the external field is assumed const. in space and time.

In truth the factor $(z-x)(1-2z+zx)$ in the numerator of eq. (45) is

$$(z-x)(1-2z+zx)=(z-x)^2(1-z)+(1-z)^2(z-x)-(z-x)^2. \tag{46}$$

Inserting (46) in (45), the first two terms of the results correspond to the logarithmic divergence term which did not contribute to the magnetic moments in the results calculated by Case. It is interesting to note that the values of the term (45) is sensitive to the parameters λ and α which have the physical meanings.

Thus, it will be conceived that the cut-off factor (31) given in § 6 plays a valuable role in the non-local meson theory. In the above calculations, we used the convergence

factor of type $\partial_c(x-\lambda^2 \cdot d^2x/d\tau^2)$ in the photon-nucleon interactions, because its term represents the anomalous magnetic moments owing to the discussions in § 6.

The complete work including this and some other problems will be published in future.

Acknowledgements

Appendix I

The Dirac wave equation according to Feynman is

$$\nabla\psi=\partial\psi/\partial\tau, \tag{A·1}$$

where he assume for $\psi(x, \tau)$ a proper time dependence of the form

$$\psi(x, \tau)=\psi(x)e^{-im\tau}. \tag{A·2}$$

The displacement of the coordinates $x_\mu(\tau)$ with respect to proper time was given by eq. (16) in this paper.

$$\left.\begin{array}{c}\sum_{n=0}^{\infty}\frac{\lambda^n}{n!}\frac{d^n x_\mu(\tau)}{d\tau^n}=x_\mu(\tau)+\lambda[\Gamma, x_\mu]+\frac{\lambda^2}{2!}[\Gamma, [\Gamma, x_\mu]]+\cdots,\\[2mm]\nabla=\gamma_\mu\partial_\mu.\end{array}\right\} \tag{A·3}$$

From eq. (A·3) we have

$$dx_\mu/d\tau \to [\nabla, x_\mu]=\gamma_\mu, \quad \text{etc.} \tag{A·4}$$

Considering the velocity $dx_\mu/d\tau$ from the classical standpoint, it has another expression

$$dx_\mu/d\tau \to \frac{x_\mu(\tau+\lambda)-x_\mu(\tau)}{\lambda}.$$

In case of the operator $Q=O(x_\mu(\tau))\cdot O(\tau)$ we obtain the form

$$\frac{dQ}{d\tau}=\sum_{\mu=1}^{4}\frac{dx_\mu}{d\tau}\frac{dO(x_\mu(\tau))}{dx_\mu}\cdot O(\tau)+O(x_\mu(\tau))\cdot\frac{dO(\tau)}{d\tau}. \tag{A·5}$$

If we take into consideration eq. (A·4) in the first term on right hand side of the above eq. (A·5), the equation (6) of this paper is obtained.

Appendix II

Equation (8) in § 2 is given by

$$\mathcal{E}(\tau-\tau')-\mathcal{E}(\tau-\tau'')=(2\pi i)^{-1}\int_{-\infty}^{+\infty}\frac{d\alpha}{\alpha}[e^{i\alpha(\tau-\tau')}-e^{-i\alpha(\tau-\tau'')}]$$

$$= \begin{cases} 0 & \text{for} \quad \tau < \tau' < \tau'' \\ 1 & \text{for} \quad \tau' < \tau < \tau'' \\ 0 & \text{for} \quad \tau' < \tau'' < \tau, \end{cases} \tag{A$\cdot$6}$$

with

$$\dot{\delta}(\tau - \tau') = \partial_\tau \varepsilon(\tau - \tau'), \quad \delta(\tau - \tau') = (2\pi)^{-1} \int_{-\infty}^{+\infty} e^{i\alpha(\tau - \tau')} d\alpha$$

and

$$\varepsilon(\tau - \tau') = (2\pi i)^{-1} \int_{-\infty}^{+\infty} \frac{d\alpha}{\alpha} e^{i\alpha(\tau - \tau')}.$$

We see that (A·6) has a value which differs from zero when τ lies inside a region (τ', τ'').

Appendix III

Equation (31) in §6 is obtained by the following method. When the factor $(2i) \cdot \sigma_{\mu\nu}(P - K)_\mu \Delta P_\nu$ is multiplied to both the numerator and the denominator of eq. (30), we get the following form as the denominator

$$4\lambda^2 [(\boldsymbol{P} - \boldsymbol{k})(\boldsymbol{\Delta P'})(\boldsymbol{P} - \boldsymbol{k})(\boldsymbol{\Delta P'}) - 2(\boldsymbol{P} - \boldsymbol{k})(\boldsymbol{\Delta P'})(P - k)_\lambda (\Delta P')_\lambda$$
$$+ \{(P - k) \cdot (\Delta P')\}^2], \tag{A$\cdot$7}$$

where $(\boldsymbol{P} - \boldsymbol{k}) = \gamma_\mu (P - k)_\mu$ and $\Delta P' = P - P'$. In this procedure, we had the use of anticomutation relations $[\gamma_\mu, \gamma_\nu] = 2\delta_{\mu\nu}$. Exchanging the positions of $\Delta \boldsymbol{P'}$ and $(\boldsymbol{P} - \boldsymbol{k})$ in the first term of the brackets in (A·7), this denominator is written over as the form

$$4\lambda^2 [\{(P - k) \cdot (\Delta P')\}^2 - (\boldsymbol{P} - \boldsymbol{k})^2 (\boldsymbol{\Delta P})^2]. \tag{A$\cdot$8}$$

A symbol (A)·(B) is a scalar product of A and B.

References

1) S. Tomonaga, Prog. Theor. Phys. **1** (1946), 27.
    J. Schwinger, Phys. Rev. **74** (1948), 1937.
    R. P. Feynman, ibid. **76** (1949), 749.
    F. Dyson, ibid. **75** (1949), 486, 1736.
2) E. C. G Stueckelberg, Phys. Rev. **81** (1951), 130.
3) E. C. G. Stueckelberg et T. A. Green, Helv. Phys. Acta **24** (1951), 153.
4) R. P. Feynman, Rev. Mod. Phys. **20** (1948), 367.
      „   „   Phys. Rev. **80** (1950), 440.
5) R. P. Feynman, Phys. Rev. **76** (1948), 769.
6) D. Bohm, M. Weinstein and H. Kouts, Phys. Rev. **76** (1949), 867.
7) G. J. Yevick, Phys. Rev. **85** (1952), 911.
8) E. C. G. Stueckelberg and A. Perturman, Phys. Rev. **82** (1951), 548.
9) K. M. Case, Phys. Rev. **76** (1948), 1.
10) J. Schwinger, Phys. Rev. **75** (1949), 651.
11) D. Bohm, Prog. Theor. Phys. **9** (1953), 273.
12) A. Pais and G. E. Uhlenbeck, Phys. Rev. **79** (1950), 145.
13) I. L. Foldy, Phys. Rev. **87** (1952), 688.
14) F. Dyson, Phys. Rev. **83** (1951), 608.
15) W. Pauli, Rev. Mod. Phys. **13** (1941), 203.

On the Theory of Cooperative Phenomena

Shigetoshi KATSURA

Department of Applied Science, Tôhoku University

(Received October 15, 1953)

The partition function of the Ising lattice which has a small number of lattice points on each side (edge) has been obtained. Such a partition function is convenient to consider several general properties of the lattice because of its closed form. The effect of the external magnetic field, Curie point of the three-dimensional lattice, and the theory of condensation by cluster integrals are discussed.

§ 1. Introduction

Statistical treatment of the Ising model is important because it is a bommon problem through ferromagnetic lattice, alloy, imperfect gas and liquid (lattice gas), etc., though it is simple for the model of the ferromagnetism. Onsager[1], Kaufman[2], Nambu[3] obtained a rigorous form of the partition function of the two-dimensional square lattice without external field, and Yang[4] obtained the spontaneous magnetization.

In the general case under arbitrary finite magnetic field, however, partition function as a closed binary function of temperature and external field has not yet been obtained, though the expansions by Domb[5] and Tanaka[6] exist. For the three-dimensional cubic lattice, Tanaka, Katsumori and Toshima (T.K.T.)[7], Oguchi[8] and Wakefield[9] obtained partition functions by expansions, but it also remains to obtain a solution in a closed form.

Meanwhile, in the theory of condensation by Mayer[10], Kahn and Uhlenbeck[11], Born and Fuchs[12], using the expansions by cluster integrals , it is doubtful whether the real positive singularity of the analytic function continued from the power series $\sum_{l=1}^{\infty} b_l z^l$, would give the point of condensation or not[13][14][15][16][17]. To investigate this problem, one must regard cluster integral as a function of the total volume of the system and examine the volume dependency of $b_l(V)$. In this paper the author treats the imperfect gas as a lattice gas (neglecting a free volume), an alloy of molecules and holes with negative interaction which can exist only on the lattice points.

The phase transition is mainly treated as a mathematical singularity of the infinitely large system. However, the phase transition is to be characterized by the fact that the partition function (or some functions derived from it) of finite system suffers a sudden change at some value of the parameter. And it is necessary and sufficient to prove that partition function of finite system approaches to the function with angle or jump as the size of the system approaches to infinity. In this paper the author calculates the partition function of the system which has finite number of lattice points as its volume. The essential

features of the cooperative phenomena can be seen even in pretty small systems as described below. The extrapolation of these results will enable us to infer a qualitative and quantitative nature of large systems.

We consider the square (cube) lattice which has a finite number of lattice points on each side (edge). The partition function of such system can be obtained in a *closed form** as a function of temperature and external field, and so the whole aspects, especially the symmetric properties of the partition function, are easily seen. Domb, T.K.T.,[7] Oguchi[8] and Wakefield[9] have obtained the approximation to the n-th power of β, while our expression may be regarded also as n-th approximation in a different sense. The aim of the present paper is, to see the feature of the partition function under arbitrary magnetic field, to obtain an approximate solution of the three-dimensional cubic Ising lattice, and to discuss the theory of condensation, by the consideration of such finite systems.

§ 2. Partition function of finite systems

The partition function of the square (cube) lattice system whose each side (edge) has m, n (l, m, n) lattice points can be obtained either by elementary enumeration of complexions or by matrix method. It is denoted by $(m \times n)$ or $(l \times m \times n)$. We represent the nearest neighbour interaction by J, spin magnetic moment by μ, external magnetic field by $\mathcal{H}$, and

$$H = J/2kT, \quad I = \mu \mathcal{H}/kT, \quad \beta = e^{-2H}, \quad \alpha = e^I. \tag{1}$$

The partition function Ξ in case of bounded finiteness is given by

$$\Xi = \sum \sum (V_3 V_2' V_1)^{m-1} V_3 V_2', \tag{2}$$

and the one in case of periodic condition is

$$\Xi = \text{trace} (V_3 V_2 V_1)^m, \tag{3}$$

where

$$V_1 = \prod_{i=1}^{n} (e^H + C_i e^{-H}), \tag{4}$$

$$V_2 = \exp\left\{ H \sum_{i=1}^{n-1} s_i s_{i+1} + H s_n s_1 \right\}, \tag{5}$$

$$V_2' = \exp\left\{ H \sum_{i=1}^{n-1} s_i s_{i+1} \right\}, \tag{6}$$

$$V_3 = \exp\left\{ I \sum_{i=1}^{n} s_i \right\},$$

$$s_i = 1 \times 1 \times \cdots \times \begin{pmatrix} 1 & \\ & -1 \end{pmatrix} \times 1 \times \cdots \times 1, \tag{7}$$

$$C_i = 1 \times 1 \times \cdots \times \begin{pmatrix} & 1 \\ 1 & \end{pmatrix} \times 1 \times \cdots \times 1.$$

---

* When we want to obtain a partition function per particle, we do not expand the fractional power $(1 + a_1 \beta + a_2 \beta^2 + \cdots)^{1/N}$ in the form of power series, but evaluate it numerically.

The relation to the case of imperfect gas is discussed by Rushbrooke[18] Tanaka[6)7], Ono[17], Yang and Lee[19].

$$\Xi_{gas}=\sum_{n=0}^{N}\Omega_{n}z^{n}=\sum_{n=0}^{N}Q_{n}\alpha^{2n}=\alpha^{N}\beta^{N}\Xi \qquad \text{for square lattice}$$

$$=\alpha^{N}\beta^{3N/2}\Xi \qquad \text{for cubic lattice.} \tag{8}$$

When we expand the left hand side of (8) by fugacity $z=\alpha^{2}\beta^{4}$ (for square lattice) or $z=\alpha^{2}\beta^{6}$ (for cube lattice), the coefficient $\Omega_{n}=Q_{n}\beta^{-4n}$ or $Q_{n}\beta^{-6n}$ of z^{n} is seen to be a partition function of the canonical ensemble, and Ξ_{gas} to be a grand partition function of the grand canonical ensemble of lattice gas systems. It is equivalent to put many finite systems into a particle reservoir, each system can interchange particles but no interaction exist between particles of different systems.

The dependency of the partition function (per particle) or of the cluster integrals on the volume of the total system is mainly due to the contribution of the closely packed state, here the boundedness of the volume by surface is only a secondary effect. Therefore the consideration of the torus of finite size is sufficient for the effect of the finiteness of the volume. We consider mainly (3) hereafter. The partition functions are as follows:

$$(2\times 2) \qquad \Xi(\alpha,\beta)=\beta^{-1}(\alpha^{4}+\alpha^{-4})+4(\alpha^{2}+\alpha^{-2})+(4+2\beta^{4}), \tag{9}$$

$$\Xi(1,\beta)=2(\beta^{4}+6+\beta^{-4}), \tag{10}$$

$$\Xi(1,\beta,\beta')=2(\beta^{2}\beta'^{2}+\beta^{-2}\beta'^{-2}+\beta^{2}\beta'^{-2}+\beta'^{-2}\beta^{2}+4), \tag{11}$$

$$(2\times 3) \qquad \Xi(1,\beta)=2(\beta^{-6}+9\beta^{-2}+13+6\beta^{2}+3\beta^{4}), \tag{12}$$

$$(3\times 3) \qquad \Xi(\alpha,\beta)=(\alpha^{9}+\alpha^{-9})\beta^{-9}+(\alpha^{7}+\alpha^{-7})9\beta^{-6}$$

$$+(\alpha^{5}+\alpha^{-5})(18\beta^{-3}+18\beta^{-1})$$

$$+(\alpha^{3}+\alpha^{-3})(6\beta^{-3}+36\beta^{-1}+36\beta+6\beta^{3})$$

$$+(\alpha+\alpha^{-1})(45\beta^{-1}+36\beta+45\beta^{3}), \tag{13}$$

$$\Xi(1,\beta)=2(\beta^{-9}+9\beta^{-5}+24\beta^{-3}+99\beta^{-1}+72\beta+51\beta^{3}). \tag{14}$$

Fig. 1 shows the behaviour of the surface $\Xi(\alpha,\beta)$ in (α,β) plane. $\Xi(\alpha)$ takes a minimum value at $\alpha=1$ when β is fixed In low temperature this minimum is a steeply sloping valley and in high temperature it is gently sloping valley. In the limit of n, $m\rightarrow\infty$, this valley becomes angle when $\beta<\beta_{c}$, while $(\partial\Xi/\partial\alpha)_{\alpha=1}=0$ when $\beta>\beta_{c}$.

Approximate singular property of $\Xi(1,\beta)$ near the Curie point $\beta\simeq\beta_{c}$ is difficult to see in this form. Similar transformation by

$$S=S^{-1}=\frac{1}{2}(1\pm s_{1})\frac{1}{\sqrt{2}}(s_{1}+U)$$
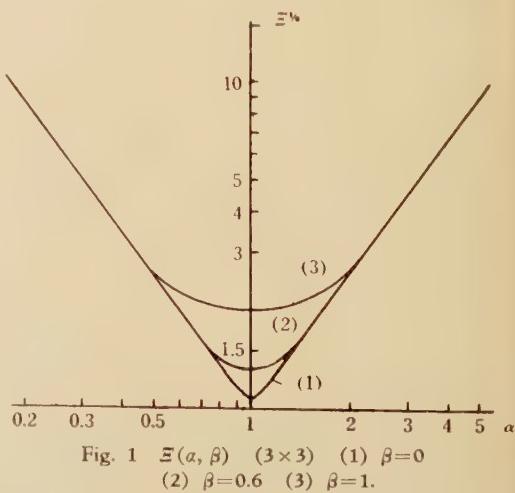
Fig. 1 $\Xi(\alpha,\beta)$ (3×3) (1) $\beta=0$
(2) $\beta=0.6$ (3) $\beta=1$.

(where $U = C_1 C_2 \cdots C_n$) dissolves $V_2 V_1$ to the operators in the even and the odd spaces. They are denoted by Ξ^+ add Ξ^- :

$$\Xi^\pm(1, \beta) = \text{trace} \frac{1}{2}(1 \pm s_1)\frac{1}{\sqrt{2}}(s_1 + U) V_2 V_1 \frac{1}{\sqrt{2}}(s_1 + U)\frac{1}{2}(1 \pm s_1). \quad (15)^*$$

In the limit $n, m \to \infty$, Ξ^+ and Ξ^- degenerate below the Curie point $\beta_c = \sqrt{2} - 1$, and separate each other for $\beta > \beta_c$. They are not branches of one analytic function. Ξ^+ and Ξ^- of finite system also *degenerate approximately* below the Curie point. Numerical calculation of $\dfrac{\Xi^+(1, \beta, \beta)^{1/(m \times n)}}{2 \text{ ch } 2H}$ and $\dfrac{\Xi^-(1, \beta, \beta)^{1/(m \times n)}}{2 \text{ ch } 2H}$ shows that the former and the latter take close values in $0 < \beta < \beta_c = \sqrt{2} - 1$, and separate each other for $\beta_c < \beta \to 1$ (Fig. 2). Particularly, the former can be expressed as a function
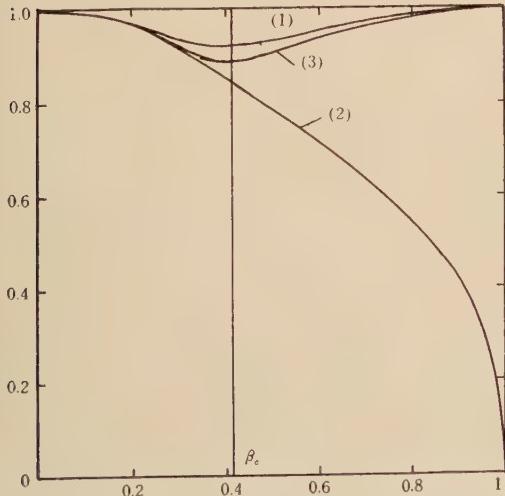
Fig. 2   (1) $\{\Xi^+_{3 \times 3}\}^{1/9}/2$ ch $2H$

        (2) $\{\Xi^-_{3 \times 3}\}^{1/9}/2$ ch $2H$

        (3) $\lambda_{\infty \times \infty}/2$ ch $2H$ (Onsager)
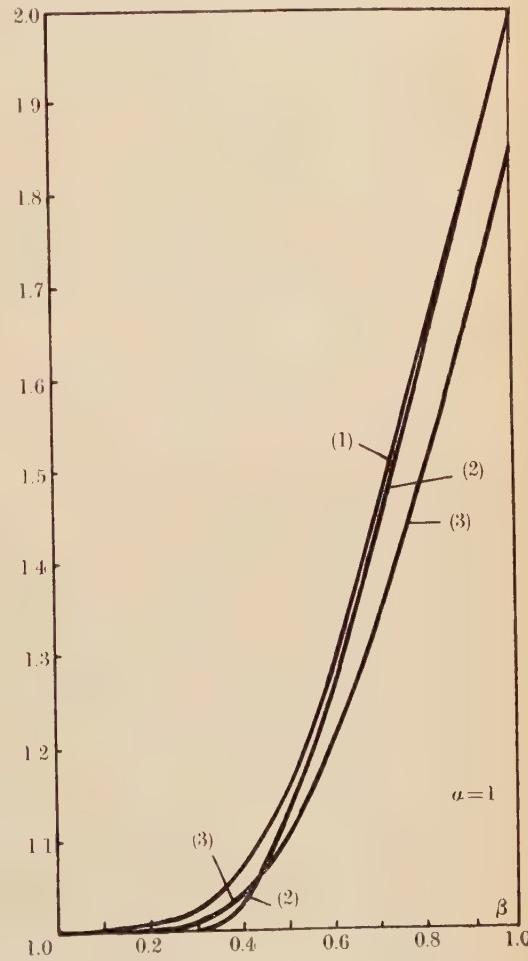
Fig. 3   (1) $\beta\{\Xi^+_{3 \times 3}\}^{1/9}$

        (2) $\beta\lambda_{\infty \times \infty}$ (Onsager)

        (3) $\beta\{\Xi_{3 \times 3}/2\}^{1/9}$.

* (3×3)

$$\Xi^\pm(1, \beta) = \beta^{-9} \pm 3\beta^{-6} + 9\beta^{-5} \pm 18\beta^{-4} + 24\beta^{-3} \pm 45\beta^{-2} + 99\beta^{-1} \pm 124$$

$$+ 72\beta \pm 45\beta^2 + 51\beta^3 \pm 18\beta^4 \pm 3\beta^6.$$

of $k = \dfrac{2\,\mathrm{sh}\,2H}{\mathrm{ch}^2\,2H} = 4\kappa$ only, and its minimum value always agrees exactly to $\beta_C = \sqrt{2}-1$ regardless of m and n. Ξ^1 can be adopted as a good approximation to the exact solution in the whole region of β (Figs. 2 and 3).

Three-dimensional case.

The partition function is given by

$$\Xi(\alpha,\beta) = \sum \sum (V_3 V_2' \tilde{V}_2' V_1)^{n-1} V_3 V_2' \tilde{V}_2', \tag{16}$$

$$V_3 = \prod_{j=1}^{l} \prod_{i=1}^{m} \exp\{l\,s_{ij}\},$$

$$V_2' = \prod_{i=1}^{l} \exp\{H(s_{i1}s_{i2} + s_{i2}s_{i3} + \cdots + s_{i,m-1}s_{im})\},$$

$$\tilde{V}_2' = \prod_{j=1}^{m} \exp\{H(s_{1j}s_{2j} + s_{2j}s_{3j} + \cdots + s_{l-1,j}s_{lj})\}, \tag{17}$$

$$V_1 = \prod_{j=1}^{m} \prod_{i=1}^{l} (e^H + C_{ij}e^{-H}).$$

In case of periodic conditions,

$$\Xi(\alpha,\beta) = \mathrm{trace}(V_3 V_2 \tilde{V}_2 V_1)^n, \tag{18}$$

$$V_2 = \prod_{i=1}^{l} \exp\{H(s_{i1}s_{i2} + s_{i2}s_{i3} + \cdots + s_{i,m-1}s_{im} + s_{im}s_{i1})\},$$

$$\tilde{V}_2 = \prod_{j=1}^{m} \exp\{H(s_{1j}s_{2j} + s_{2j}s_{3j} + \cdots + s_{l-1,j}s_{lj} + s_{lj}s_{1j})\}, \tag{19}$$

where

$$s_{ij} = 1 \times 1 \times \cdots \times \begin{pmatrix} 1 & \\ & -1 \end{pmatrix} \times 1 \times \cdots \times 1,$$

$$C_{ij} = 1 \times 1 \times \cdots \times \begin{pmatrix} & 1 \\ 1 & \end{pmatrix} \times 1 \times \cdots \times 1. \tag{20}$$

$$\underbrace{\phantom{C_{ij} = 1 \times 1 \times \cdots}}_{(j-1)l+i}$$

The results are as follows* :

$(2 \times 2 \times 2)$

$$\Xi(\alpha,\beta) = \beta^{-12}(\alpha^8 + \alpha^{-8}) + 8\beta^{-6}(\alpha^6 + \alpha^{-6}) + (12\beta^{-4} + 16)(\alpha^4 + \alpha^{-4})$$
$$+ (24\beta^{-2} + 24\beta^2 + 8\beta^6)(\alpha^2 + \alpha^{-2}) + (6\beta^{-4} + 32 + 30\beta^4 + 2\beta^{12}), \tag{21}$$

$$\Xi(1,\beta) = 2(\beta^{-12} + 8\beta^{-6} + 15\beta^{-4} + 24\beta^{-2} + 32 + 24\beta^2 + 15\beta^4 + 8\beta^6 + \beta^{12}). \tag{22}$$

$(3 \times 3 \times 2)**$

$$\Xi(\alpha,\beta) = \beta^{-27}(\alpha^{18} + \alpha^{-19}) + 18\beta^{-21}(\alpha^{16} + \alpha^{-16})$$
$$+ (9\beta^{-19} + 36\beta^{-17} + 108\beta^{-15})(\alpha^{14} + \alpha^{-14})$$
$$+ (84\beta^{-15} + 144\beta^{-13} + 324\beta^{-11} + 264\beta^{-9})(\alpha^{12} + \alpha^{-12})$$

---

* Three-dimensional torus $(2 \times 2 \times n)$ can be developed on the two-dimensional torus. Therefore we must consider at least $(3 \times 3 \times 2)$ to see the essential feature of the three-dimensional lattice. $\Xi_{3 \times 3 \times 2}$ is indebted to C. Katsura.

$$+ (18\beta^{-15} + 36\beta^{-13} + 432\beta^{-11} + 648\beta^{-9} + 864\beta^{-7} + 792\beta^{-5} + 270\beta^{-3})(\alpha^{10} + \alpha^{-10})$$

$$+ (36\beta^{-13} + 144\beta^{-11} + 666\beta^{-9} + 1224\beta^{-7} + 2484\beta^{-5} + 1728\beta^{-3} + 1530\beta^{-1}$$
$$+ 648\beta + 108\beta^3)(\alpha^8 + \alpha^{-8})$$

$$+ (6\beta^{-15} + 36\beta^{-11} + 516\beta^{-9} + 1296\beta^{-7} + 2268\beta^{-5} + 4296\beta^{-3}$$
$$+ 4428\beta^{-1} + 3168\beta + 1440\beta^3 + 954\beta^5 + 144\beta^7 + 12\beta^9)(\alpha^6 + \alpha^{-6})$$

$$+ (72\beta^{-11} + 216\beta^{-9} + 1008\beta^{-7} + 2772\beta^{-5} + 4608\beta^{-3} + 6480\beta^{-1} + 7164\beta$$
$$+ 5184\beta^3 + 2484\beta^5 + 1152\beta^7 + 576\beta^9 + 108\beta^{11})(\alpha^4 + \alpha^{-4})$$

$$+ (45\beta^{-11} + 108\beta^{-9} + 1134\beta^{-7} + 2277\beta^{-5}$$
$$+ 4833\beta^{-3} + 7992\beta^{-1} + 10440\beta$$
$$+ 7236\beta^3 + 5148\beta^5 + 2880\beta^7 + 1089\beta^9$$
$$+ 360\beta^{11} + 216\beta^{13})(\alpha^2 + \alpha^{-2})$$

$$+ (314\beta^{-9} + 576\beta^{-7} + 2754\beta^{-5}$$
$$+ 4464\beta^{-3} + 7830\beta^{-1} + 9216\beta$$
$$+ 10872\beta^3 + 6480\beta^5 + 3618\beta^7$$
$$+ 1458\beta^9 + 936\beta^{11} + 102\beta^{13}), \qquad (23)$$

$$\Xi(1,\beta) = 2(\beta^{-27} + 18\beta^{-21} + 9\beta^{-19} + 36\beta^{-17}$$
$$+ 216\beta^{-15} + 216\beta^{-13} + 1053\beta^{-11}$$
$$+ 2575\beta^{-9} + 5814\beta^{-7} + 11970\beta^{-5}$$
$$+ 17967\beta^{-3} + 24345\beta^{-1} + 26028\beta$$
$$+ 19404\beta^3 + 11826\beta^5 + 5985\beta^7 + 2406\beta^9$$
$$+ 936\beta^{11} + 216\beta^{13} + 51\beta^{15}). \qquad (24)$$

The behaviour of $\Xi(1,\beta)$ is shown in Fig. 4 with the results of Wakefield[9]. General qualitative feature is similar to the two-dimensional case.

Next, let us consider the thermodynamic quantities.*

Free energy, $F = (J/\log\beta)\log\Xi$, (25)

energy, $\qquad E = J\beta\dfrac{\partial\log\Xi}{\partial\beta} + \dfrac{J^2}{\mu\mathcal{H}}$
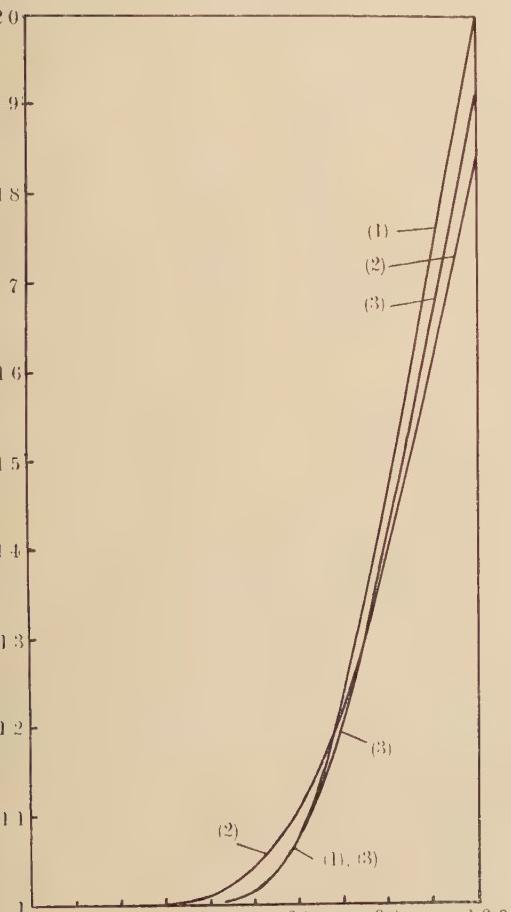
Fig. 4   (1) $\beta^{3/2}\lambda_{\infty\times\infty\times\infty}$ Wakefield
        (2) $\beta^{3/2}\{\Xi_{2\times2\times2}/2\}^{1/8}$
        (3) $\beta^{3/2}\{\Xi_{3\times3\times2}/2\}^{1/18}$

---

* In this paper we adopt the formula of the statistical mechanics which is admitted as the asymptotic form for large N. It may be allowable except the problem of the fluctuation.

$$\times \left(\frac{\log \alpha}{\log \beta} \right)^2 \alpha^2 \frac{\partial \log \Xi}{\partial (\alpha^2)}, \qquad (26)$$

specific heat,

$$C = \frac{k}{J} (\log \beta)^2 \beta \frac{\partial E}{\partial \beta}$$
$$+ \frac{k}{\mu \mathcal{H}} (\log \alpha)^2 \alpha^2 \frac{\partial E}{\partial (\alpha^2)}. \qquad (27)$$

The results of $(m \times n)$ are shown in Figs. 5 and 6 with the exact solution of Onsager[1]. The specific heat has a maximum
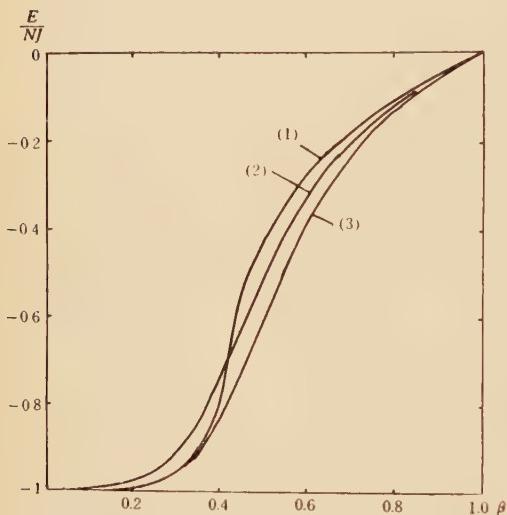

Fig. 5 E/NJ (1) $(E/NJ)_{\infty \times \infty}$
                  (2) $(E^+/NJ)_{3 \times 3}$
                  (3) $(E/NJ)_{3 \times 3}$
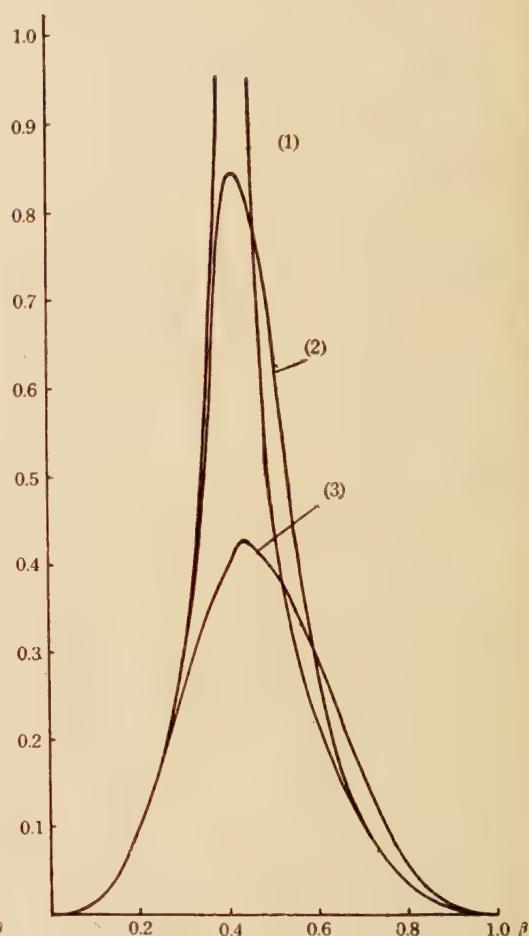
Fig. 6 C/Nk (1) $(\infty \times \infty)$ (Onsager)
                  (2) (3×3)
                  (3) (2×2)

nearly at $\beta \simeq 0.42$ when $\alpha = 1$ (without external field). Though the maximum value of the specific heat is of course finite, the whole shape of the curve is similar to that of the exact solution.

The energy and the specific that of the three-dimensional cubic lattice is shown in Figs. 7 and 8. The maximum of the specific heat lies nearly at $\beta \simeq 0.57 \sim 0.58$*. These values $\beta_c = 0.56 \sim 0.58$ may be considered as an approximation of the Curie

* To try to determine the Curie point by some kind of symmetry, we insert $V_1 V_2$ between $S = S^{-1}$ $= \frac{1}{2} (1 \pm s_{11}) \frac{1}{V} (s_{11} + U)$ similarly to the two-dimensional case. $\dfrac{\Xi^+(1, \beta)}{(2ch3H)^{l \times m \times n}}$ derived from its shows its minimum at $\beta \doteq 0.56$ for $(2 \times 2 \times 2)$ and $(2 \times 2 \times 3)$ (Fig. 9).

$(2 \times 2 \times 2)$ $\Xi^+(1, \beta) = \beta^{-12} + 2\beta^{-8} + 8\beta^{-6} + 39\beta^{-4} + 24\beta^{-2} + 108 + 24\beta^2 + 39\beta^4 + 8\beta^6 + 2\beta^8 + \beta^{12}$
            $\Xi^-(1, \beta) = \beta^{-12} - 2\beta^{-8} + 8\beta^{-6} - 9\beta^{-4} + 24\beta^{-2} - 44 + 24\beta^2 - 8\beta^4 + 8\beta^6 - 2\beta^8 + \beta^{12}$

$(2 \times 2 \times 3)$ $\Xi^+(1, \beta) = \beta^{-18} + 3\beta^{-14} + 12\beta^{-12} + 51\beta^{-10} + 36\beta^{-8} + 283\beta^{-6} + 468\beta^{-4} + 750\beta^{-2}$
                   $+ 888 + 750\beta^2 + 468\beta^4 + 283\beta^6 + 36\beta^8 + 51\beta^{10} + 12\beta^{12} + 3\beta^{14} + \beta^{18},$
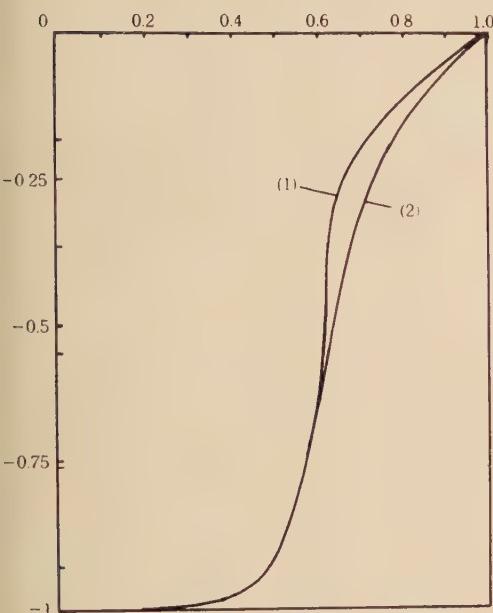
Fig. 7 E/NJ
(1) $(\infty \times \infty \times \infty)$ Wakefield
(2) $(3 \times 3 \times 2)$
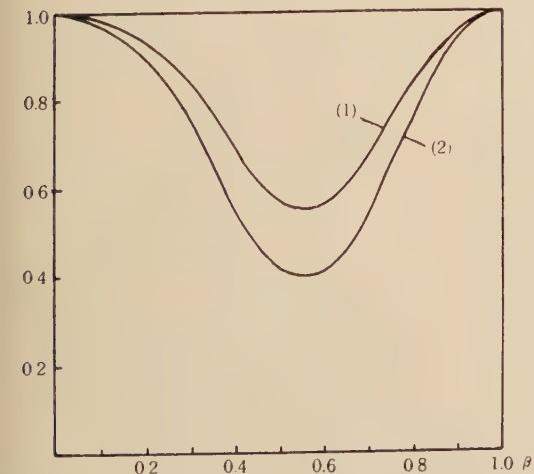
Fig. 9 $\mathcal{Z}^{+}/(2 \operatorname{ch} 3H)^{l \times m \times n}$
(1) $2 \times 2 \times 2$
(2) $2 \times 2 \times 3$
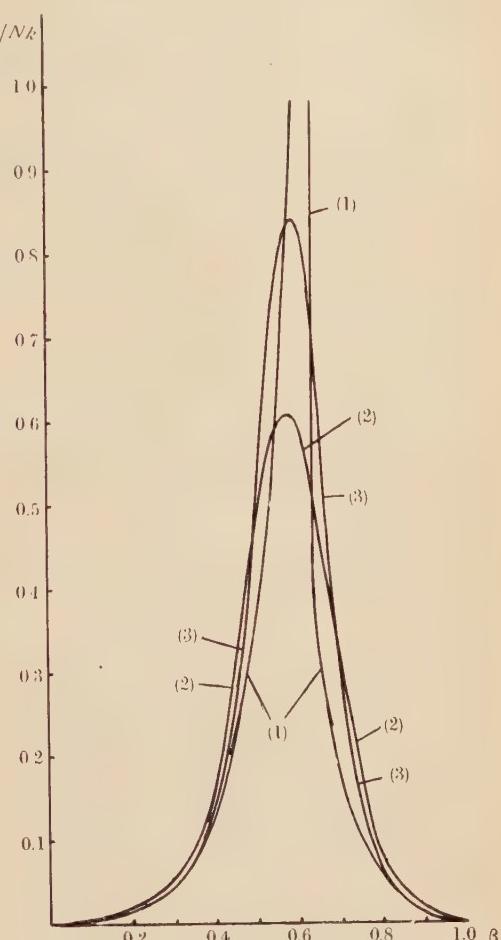
Fig. 8 C/Nk
(1) $(\infty \times \infty \times \infty)$ Wakefield
(2) $(2 \times 2 \times 2)$
(3) $(3 \times 3 \times 2)$

point of the three dimensional cubic lattice. They are somewhat lower than the values of both Oguchi[8] $0.62 < \beta_C < 0.66$ and of Wakefield[9] $\beta_C = 0.641$, while it is nearly equal to the value of plane triangular lattice $\beta_C = \dfrac{1}{\sqrt{3}} = 0.577$, Such a weak cooperative

nature in our results is interesting.   Oguchi[8] showed the similar results, that is, the correlation of simple cubic lattice is weaker than the one of plane triangular lattice.

The coffiecient of β^n in $\mathcal{Z}(\alpha, \beta)$ as a function of n shows a distribution of the energy "eigenvalues" (differential of the phase volume). We see that this function can

be well approximated by Gaussian distribution (Fig. 10). The relation to the Heisenberg theory of ferromagnetism will be discussed in another place.

Let us denote the probability that n spins in N orient to $+$ ($N-n$ spins orient to $-$) by $P_n(\alpha, \beta, N)$, and denote the value of n corresponding to the maximum probability by n'. Then the magnetization $M(\alpha, \beta)$ is given by

$$M = NJ - 2n'(\alpha, \beta)J. \quad (28)$$

The spontaneous magnetization $M(1, \beta)$ has "Curie point" $\beta'_C = 0.65$ (3×3). (Cf. Fig. 11 in the next section). The magnetization M given by

$$M = J\alpha^2 \cdot \partial \log \Xi / \partial (\alpha^2) \quad (30)$$

takes minimum at $\alpha = 1$ ($\beta = $ const.) regardless of $\beta \lessgtr \beta_C$, finite value of the spontaneous magnetization cannot be obtained from (29).

Or, we remove the terms of negative powers of (29), we get for example for (3×3),

$$M = (1 + 7_i\beta^4 + 12_i\beta^6 + 27_i\beta^8$$
$$+ 16_i\beta^{10} + 7_i\beta^{12})/(1 + 9_i\beta^4$$
$$+ 24_i\beta^6 + 99_i\beta^8 + 72_i\beta^{10}$$
$$+ 51_i\beta^{12}). \quad (30)$$

The value from (30) remains finite at $\beta = 0.41$ and the point of inflection is $\beta''_C = 0.52$. (30) has a significance only in $\beta < \beta_C$.

Fig. 10   Distribution of energy "eigenvalues". ($3 \times 3 \times 2$) The ordinate represent the coefficient of $a^m \beta^n$. The abscissa represent n. (1) The coefficient of $\Xi(1, \beta)$. (2) $m = 0$. (3) $m = \pm 2$. (4) $m = \pm 4$. (5) $m = \pm 6$. (6) $m = \pm 8$. (7) $m = \pm 10$. (8) $m = \pm 12$. (9) $m = \pm 14$. (10) $m = \pm 16$. (11) $m = \pm 18$.

§ 3.   Discussions on the theory of condensation

Now we consider the lattice gas as a model of the real imperfect gas. The configurational partition function

β=0.1

0.2

0.3

0.4

0.5

0.6

0.7

0.8

0.9

1.0

0 1 2 3 4 5 6 7 8 9 *n*   0 1 2 3 4 5 6 7 8 9 *n*
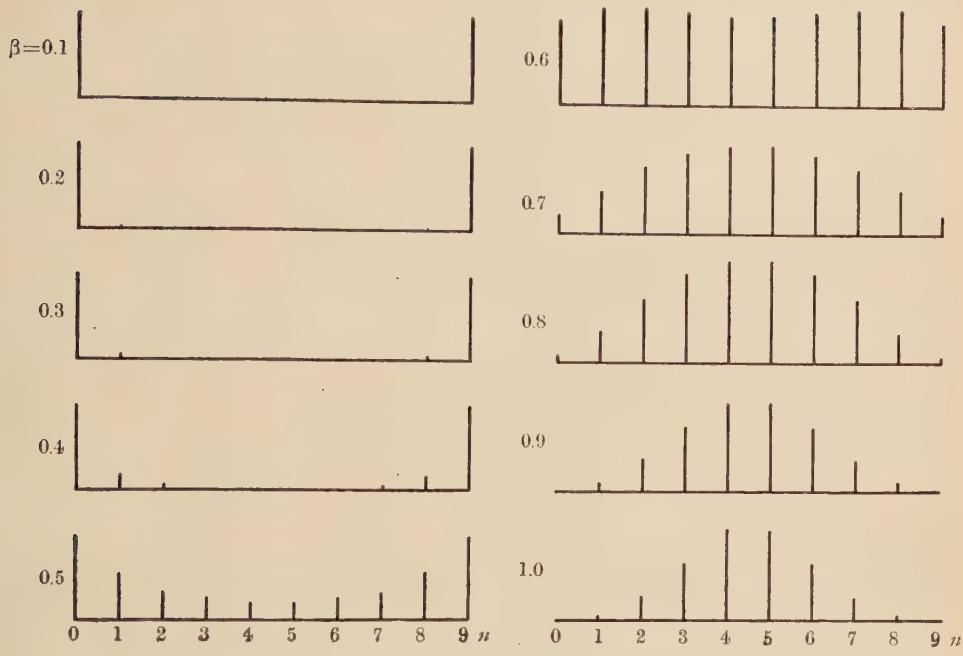
Fig. 11 $P_n(1, \beta, 3 \times 3)$

Fig. 12 $P_n(1, \beta, 3 \times 3 \times 2)$ →

$\Omega_n(\beta, N) = Q_n \beta^{-4n}$ of the system which contains n particles in N sites, and the grand partition function $\Xi_{gas}(\alpha, \beta, N)$ relates to the partition function of the Ising lattice by (8). The pressure p and the specific volume v are given by

$$pN/kT = \log \Xi_{gas} = \log \sum_{n=0}^{N} Q_n \alpha^{2n}, \quad (31a)$$

$$\frac{1}{v} = \alpha^2 \frac{\partial}{\partial(\alpha^2)} \log \Xi_{gas}$$

$$= \frac{\sum_{n=0}^{N} n Q_n \alpha^{2n}}{\sum_{n=0}^{N} Q_n \alpha^{2n}}. \quad (31b)$$

in the grand canonical ensemble. $Q_n(\beta, N) \alpha^{2n} \equiv P_n(\alpha, \beta, N)$ is a probability of finding a system which contains n particles in N sites under given temperature and fugacity (Figs. 11 and 12). This probability regarded as a function of n, has one maximum in the region of high temperature and

$\beta^2=0.1$

$\beta^2=0.2$

$\beta^2=0.3$

$\beta^2=0.4$

$\beta^2=0.5$

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 *n*
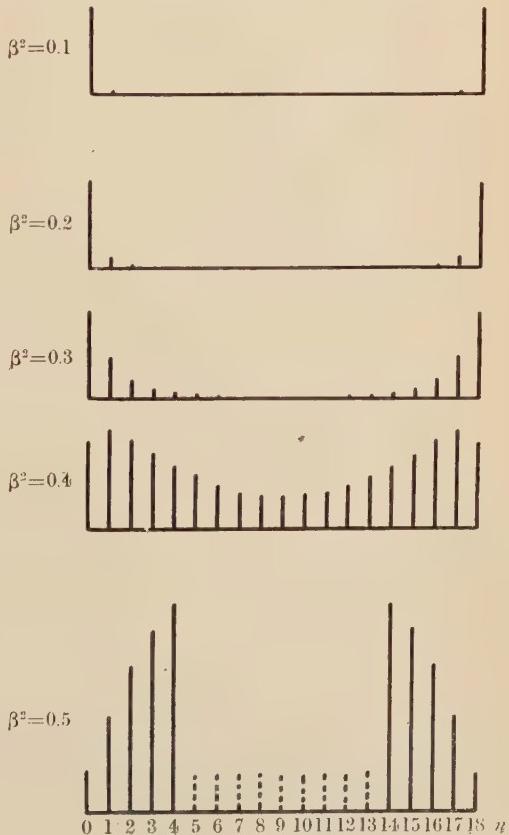
two maxima in the region of low temperature. Denoting the positions of the maxima in low temperature region by n_g and n_l $(=N-n_g, n_l > n_g)$, we see that :

$$\sum_{n \sim n_g} P_n(\alpha, \beta, N) \gg \sum_{n \sim n_l} P_n(\alpha, \beta, N) \quad \text{when } \alpha < 1,$$
$$\sum_{n \sim n_g} P_n(\alpha, \beta, N) = \sum_{n \sim n_l} P_n(\alpha, \beta, N) \qquad \alpha = 1, \qquad (32)$$
$$\sum_{n \sim n_g} P_n(\alpha, \beta, N) \ll \sum_{n \sim n_l} P_n(\alpha, \beta, N) \qquad \alpha > 1.$$

Therefore

$$1/v \doteq n_g/N \quad \text{for} \quad \alpha < 1, \quad \text{and} \quad 1/v \doteq n_l/N \quad \text{for} \quad \alpha > 1. \qquad (33)$$

$1/v$ is a one-valued monotonous increasing function of α. It changes from n_0/N to n_l/N suddenly near $\alpha \simeq 1$ and does not contain the unstable state (Fig. 13 (1)). In this way we can draw the phase boundary curve in (α, β) plane. Index g corresponds to the gaseous phase and l to the liquid phase. On the line $\alpha = 1$, two maxima merge into one at some value of $\beta = \beta_c'$. The author and Fujita have discussed such features and characterized the condensation by the fact that two phases coexist which are distinguished by density in the previous paper[20]. Such features can be seen even in small systems described here. The probability $P_n(\alpha, \beta, N)$ becomes sharp when the number of sites increases.

When some inhibition disturbs the appearance of the second maximum P_{nl} in the way of α from 0 to ∞, the curve $v(\alpha)$ does not jump at $\alpha \simeq 1$ but can be extended to the region $\alpha > 1$. This is the state of supersaturation[20]. This state is considered to be given by

$$pN/kT = \log \sum_{n=1}^{N} Q_n' \alpha^{2n}, \quad (34a)$$

$$\frac{1}{v} = \frac{\sum_{i=0}^{N} n Q_n' \alpha^{2n}}{\sum_{n=0}^{N} Q_n' \alpha^{2n}}, \qquad (34b)$$

where Q_n' is the extrapolated value of Q_n after whose second maximum is excluded. In sufficiently low temperature $P_n(1, \beta, N)$ may be put to be equal to zero for $n' < n < N$, where

Fig. 13 $1/v$ versus α (3×3)

$$(1) \quad \frac{\sum_{n=0}^{N} n Q_n \alpha^{2n}}{\sum_{n=0}^{N} Q_n \alpha^{2n}} \qquad (2) \quad \sum_{l=1}^{\infty} b_l(N) \alpha^{2l}$$

$$(3) \quad \frac{\sum_{n=0}^{N} n Q_n' \alpha^{2n}}{\sum_{n=0}^{N} Q_n' \alpha^{2n}} \qquad \sum_{l=1}^{\infty} b_l'(N) \alpha^{2l} \simeq \sum_{l=1}^{\infty} l_l(\infty) \alpha^{2l}$$

n' is the value of n corresponding to the minimum value of $P_n(1, \beta, N)$. The behaviour of $v(\alpha)$ in (34) is shown in Fig. 13 (3).

Q_n in (31) can be expressed in terms of Mayer's cluster integrals b_l, as

$$Q_n(\beta, N) = \sum_{\Sigma i n_i = n} \cdot \Pi \frac{(N b_i(\beta, N))^{n_i}}{n_i!}, \tag{35}$$

In the lattice gas model, b_l is expressed by multi-summation,

$$b_l(\beta, N) = \frac{\beta^{4l}}{l! \, m^l \, n^l} \sum \Pi f_{ij}. \tag{36}$$

This summation is taken over the number of manners of arranging l particles in $m \times n = N$ sites *allowing the multiple occupation,** and so b_l exist for all values of l ($1 \leq l < \infty$) even if N is finite. f_{ij} is

$$f_{ij} = -1 \qquad \text{when} \quad r_i = r_j,$$

$$f_{ij} = \beta^{-2} - 1 \quad \text{when} \quad r_i = r_j + (1.0) \text{ or } (0.1), \tag{37}$$

$$f_{ij} = 0 \qquad \text{otherwise.}$$

Mayer[10] and others[11][12] discussed and concluded that the point of condensation would be given by the smallest real positive singularity of an analytic function continued analytically from the power series

$$p/kT = \sum_{l=1}^{\infty} \lim_{N \to \infty} b_l(N) \alpha^{2l}, \tag{38a}$$

$$1/v = \sum_{l=1}^{\infty} l \cdot \lim_{N \to \infty} b_l(N) \alpha^{2l}. \tag{38b}$$

However, they should be replaced by

$$p/kT = \lim_{N \to \infty} \sum_{l=1}^{\infty} b_l(N) \alpha^{2l}, \tag{39a}$$

$$1/v = \lim_{N \to \infty} \sum_{l=1}^{\infty} l \, b_l(N) \alpha^{2l}. \tag{39b}$$

The author and Fujita have discussed in the previous paper[13] that (38) and (39) give different function in general. Ono[14] discussed on this problem by the method of crystal statistics. Ikeda[16] criticised Ono's discussions.[14][17] Yang and Lee[19] proved that the singularity of $\sum_{l=1}^{\infty} b_l(N) \alpha^{2l}$ are determined by the zeros of $\sum_{n=1}^{N} Q_n(\beta, N) \alpha^{2n}$, which lie on the unit circle of $|\alpha| = 1$, and gave the criterion of the condensation by the fact that the distribution of these zeros in the complex α plane approached to the positive real axis when

---

* This expression of b_l is equal to the expression

$$b_l(N, \beta) = \frac{1}{N} \sum_{\Sigma i n_i = l} (-)^{\Sigma n_i - 1} (\Sigma n_i - 1)! \, \Pi_{i=1} \frac{(Q_i(\beta, N))^{n_i}}{n_i!},$$

(S. Ono) which is obtained as the expansion coefficient of α^{2l} in

$$\log \left(\sum_{n=1}^{N} Q_n \alpha^{2n} \right) = \sum_{l=1}^{\infty} b_l \, \alpha^{2l}.$$

N approached to infinity. Yang and Lee, however, used the same symbol t_1 for two different conception.* The one is the limiting point of the zeros (we denote it z_{Yang}) and the other is z_s. They have not proved $z_{COND} \to z_s$ but only proved $z_{COND} \to z_{Yang}$ ($=z*$). The problem whether the real positive analytic singularity z_s of (38) and the irregularity $z*$ of the limit function (39) coincides or not, is not yet solved. So let us discuss on this problem. First, the volume dependency of b_l is considered.

Fujita and the author have already reported the volume dependency of the cluster integrals of the lattice gas taking the boundary surface effect into consideration.[15] The volume dependency is, however, independent of the boundary effect as described above, and it is due to the contribution of the closely packed states. Therefore we regard our system not as the infinite system under periodic condition, but as the finite system spread on the finite torus. The zeros of $\sum_{n=0}^{N} Q_n \alpha^{2n}$ are distributed on the circle of $|\alpha|=1$. If we denote the argument of the zeros by θ_n, the cluster integral b_l is expressed by[10]

$$b_l(N) = -\frac{1}{l} \frac{1}{N} \sum_{n=1}^{N} \cos l\theta_n(N). \qquad (40)**$$

The behaviour of b_l in (3×3) where $\beta^2 = 0.1$ is shown in Fig. 14. The behaviour of $b_l(N)$ regarded as a function of l is not a monotonous function: in low temperatures b_l decreases with l first, then increases and takes a large value at $l \lesssim N$, takes negative value in $l \gtrsim N$ and approaches to zero for $l \to \infty$ from negative side. Such a feature of b_l generally agrees to the supposition proposed by the author in the previous paper[13]. As the origin of the singular property of $b_l(N)$ at $l \sim N$ is due to the second maximum of $P_n(\beta, N, 1) = Q_n(\beta, N)$, such volume dependency of b_l on N makes us suppose that the difference arises between $\lim_{N \to \infty} \sum_{l=1}^{\infty} b_l(N) \alpha^{2l}$ and $\sum_{l=1}^{\infty} \lim_{N \to \infty} b_l(N) \alpha^{2l}$. That is, the contribution of the second maximum of $P_n(1, \beta, N)$ will be excluded in $\sum_{l=1}^{\infty} \lim_{N \to \infty} b_l(N) \alpha^{2l}$. If we remove this effect, $b_l(N)$ is considered to decrease monotonously with l and is regarded to be equal to $b_l(\infty)$. The numerical calculation of

$$\frac{p}{kT} = \sum b_l(\infty) \alpha^{2l} \doteqdot \sum_{l=1}^{\infty} b_l'(N) \alpha^{2l}, \qquad (41a)$$

$$\frac{1}{v} = \sum_{l=1}^{\infty} l b_l(\infty) \alpha^{2l} \doteqdot \sum l b_l'(N) \alpha^{2l}, \qquad (41b)$$

(where b_l' is extrapolated value of b_l after excluding the singular part of b_l at $N \sim l$) shows that the steep change at $\alpha \sim 1$ of $v(\alpha)$ disappears and that $v(\alpha)$ by (41) can be extended into the region $\alpha \, 1$ (Fig. 13 (3)). This curve (41) agrees with (34) farely

---

* In this point the author is indebted to Mr. K. Ikeda.

** $256 v^4 + (576\beta^4 - 576) v^3 + (288\beta^8 + 288\beta^6 - 1008\beta^4 + 432) v^2$
$\quad + (-120 + 504\beta^4 - 336\beta^6 - 216\beta^8 + 144\beta^{10} + 24\beta^{12}) v$
$\quad + (9 - 63\beta^4 + 72\beta^6 + 27\beta^8 - 72\beta^{10} + 27\beta^{12}) = 0,$ (for (3×3)).
$\qquad v = \cos^2(\theta/2), \qquad \theta_0 = \pi, \qquad \alpha = e^{i(\theta/2)},$

well. On the other hand, the calculation of $\sum_{l=1}^{\infty} l\, b_l(N)\alpha^{2l}$ shows the steep ascent nearly at $\alpha \simeq 1$. In this case $|\alpha|=1$ is the radius of convergence, so the calculated value does not extend to the region $\alpha > 1$, but its analytic continuation is equal to (31), which has no real positive singularity and whose part for $\alpha > 1$ expresses the equation of state of the liquid phase. Now, the singularity of (41) is considered to be given by zeros of $\sum_{n=0}^{N} Q_n{}' \alpha^{2n} \simeq \sum_{n=0}^{N/2} Q_n \alpha^{2n}$, which does not lie on the unit circle, and is considered to correspond to the end point of the supersaturation.

Where shall the singularity of $\sum_{l=1}^{\infty} \lim_{N\to\infty} b_l(N)\alpha^{2l}$ appear? If $b_l(N)$ is always positive and increasing monotonously with N, then $z^*=z_s$.[13] The situation, however, seems to the author not so simple. It has commonly been said that "at sufficiently low temperature $b_l(\infty)$ would be always positive." This expression should be interpreted, however, that "$b(\infty)=\lim_{N\to\infty} b_l(T, N)$ becomes at last positive for fixed l if we make T sufficiently low," and may not necessarily mean that "$b_l(\infty)$ for all values of l is positive under fixed small T." These features are supposed by the temperature dependency of $b_l(\infty \times \infty)$[15] for small l.

$$b_1(\infty \times \infty)=\beta^4,$$

$$b_2(\infty \times \infty)=\beta^6\left(2-\frac{5}{2}\beta^2\right),$$

$$b_3(\infty \times \infty)=\beta^8\left(6-16\beta^2+\frac{31}{3}\beta^4\right), \tag{42}$$

$$b_4(\infty \times \infty)=\beta^8\left(1+18\beta^2-85\beta^4+118\beta^6-\frac{209}{4}\beta^8\right),$$

$$b_5(\infty \times \infty)=\beta^{10}\left(8+43\beta^2-400\beta^4+92\beta^6-872\beta^8+\frac{1476}{5}\beta^{10}\right).$$
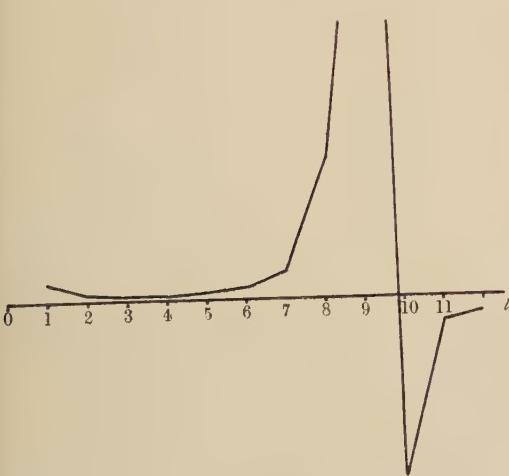
Fig. 14 $9l\,b_l\,(3\times3)$ $\beta^2=0.1$

Such temperature dependency is shown in Fig. 15.

Next let us investigate how the approximate degeneracy of the eigenvalues of $V_3 V_2 V_1$ is removed under the perturbation of the external magnetic field. In (2×2)-system the eigenvalues λ are given by

$$\lambda_1=1-\beta^2$$

$$\lambda^3-\lambda^2\{1+\beta^2+\beta^{-2}(\alpha^2+\alpha^{-2})\}$$

$$+\lambda\{(\alpha^2+\alpha^{-2}+1+\beta^{-2})(\beta^{-2}-1)\}$$

$$-(\beta^{-2}-1)(\beta^2+\beta^{-2}-2)^2=0. \tag{43}$$

S. Katsura

Fig. 15 $b_l (\infty \times \infty)$ versus β

Fig. 16(a) λ (eigenvalue of $V_3 V_2 V_1$) versus a.

$\beta = 0.1$ $e^I = a$. $I = \mu \mathcal{H} / kT$ (2×2)

Fig. 16(b) λ (eigenvalue of $V_3 V_2 V_1$) versus a

$\beta = 0.9$ $e^I = a$. $I = \mu \mathcal{H} / kT$ (2×2)
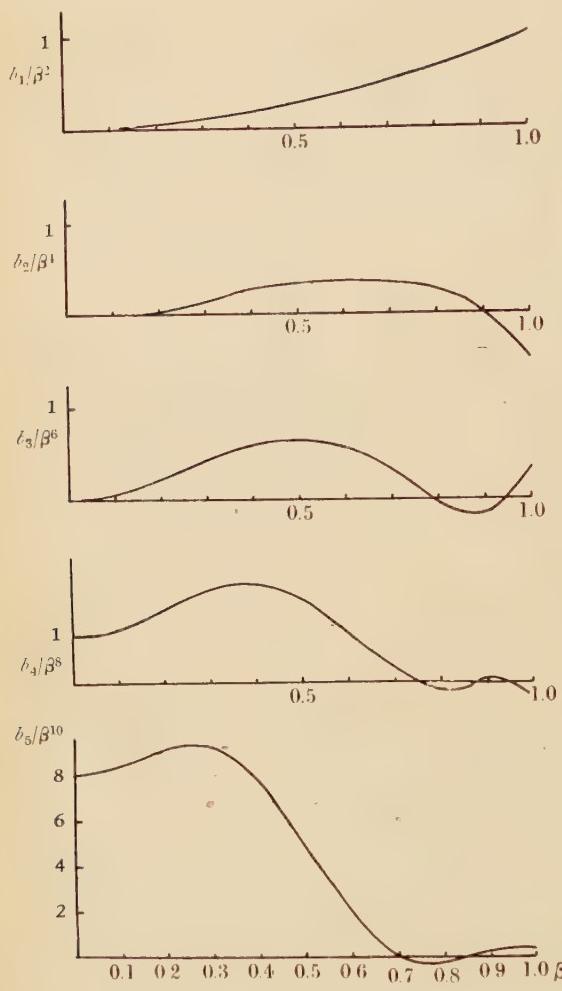
Fig. 16 shows λ as a function of α. At $\alpha = 1$, λ_1^+ and λ_1^- are perfectly separated for $\beta > \beta_c$ and nearly degenerated for $\beta < \beta_c$. This approximate degeneracy is removed by the perturbation of the external magnetic field. In this case λ_1^+ for $\alpha > 1$ is almost an extension of λ_1^- for $\alpha < 1$ and vice versa. The angle of λ_1^+, λ_1^- at $\alpha = 1$ becomes steeper as $N \to \infty$. The behaviour of the analytic continuation of the limit function of the hyperbola makes us suppose that the analytic continuation of λ_1^+ of $(\infty \times \infty)$ for $\alpha < 1$ has no singularity at $\alpha = 1$, and will be connected smoothly to λ_1^- for $\alpha > 1$. The behaviour of the limit function near $\alpha = 1$ will be analogous to that of

$$\lambda^2 \{\lambda + (\beta - \beta_c)\} = (\log \alpha)^2 \tag{44}$$

or so.
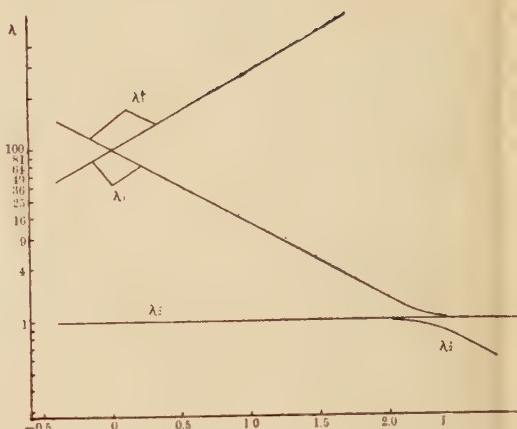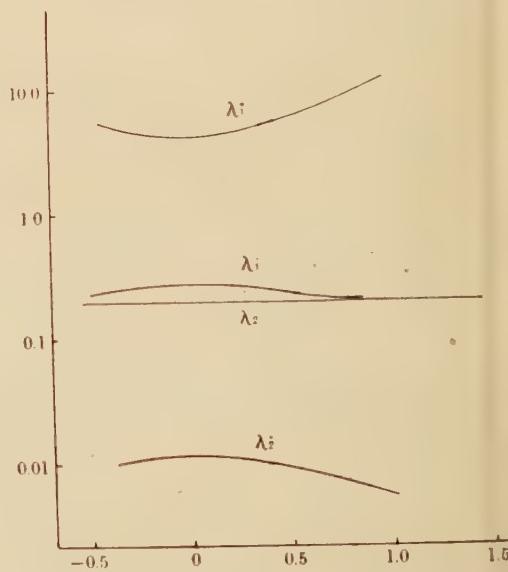
As described above the nature of the point of condensation z_{COND} is clarified as follows: z_{COND} is a point of z where the gaseous phase and liquid phase reach equilibrium in the grand canonical ensemble* $(\sum_{n \sim n_g} P_n(z, \beta, N) = \sum_{n \sim n_l} P_n(z, \beta, N))$ and hence $1/v$

$$= (\sum_{n=0}^{N} n \Omega_n(N) z^n / \sum_{n=0}^{N} \Omega_n(N) z^n) = \sum_{l=1}^{\infty} l b_l(N) z^l$$ suffers a steep ascent for finite N.

z_{COND} becomes to the irregular point z^* of the limit function $\lim_{N \to \infty} (\sum_{n=0}^{N} n \Omega_n(N) z_n / \sum_{n=0}^{N} \Omega_n(N) z^n)$

$== \lim_{N \to \infty} \sum_{l=1}^{\infty} l b_l(N) z^l$. Though it cannot be declared where shall the analytic singularity

z_s of $\sum_{l=1}^{\infty} \lim_{N \to \infty} b_l(N) z^l$ appear, it is of no importance whether z_s coincides to z^* or not. Because the physical infinity should always be the limit of the finiteness, and the limiting process in physics always requires the uniform convergence.

If one would reinterpret the conclusion of the theory by cluster integrals by the fact that the radius of convergence would give the condensation point, it is not adequate because there is no difference between above and below the critical temperature.

The origin of the condensation is the equilibrium between the gaseous phase and the condensed phase. The imprudent interchange of the two limit process $N \to \infty$ and $\sum_{l=1}^{\infty}$ (in the meaning of analytic continuation) will make us overlook this symmetric property and push the condensed phase infinitely far away. Such symmetric properties are clarified by the correct limit process and the consideration of the volume dependency of the partition function and cluster integrals.

§ 4. Conclusion

We have obtained the qualitative and quantitative information of the real lattice system by considering the comparatively small system.

The separated partition function Ξ^+ of small system is a good approximation to the partition function of the large system through the whole region of temperature, both in low and high temperature region.

The position of the maximum of the specific heat of (2×2) and (3×3) is $\beta_c = 0.42$, nearly equal to the exact value. That of $(2 \times 2 \times 2)$ and $(2 \times 3 \times 3)$ is $\beta_c = 0.57 \sim 0.58$, which is somewhat lower than the values of Curie point calculated by Oguchi[3] and Wakefield[1], and nearly equal to that of plane triangular lattice.

The distribution of the energy "eigenvalues" is well approximated by Gaussian functions. The probability $P_n(\alpha, \beta, N)$, n spins in N orient to $+$, has two maxima in low temperature region and these two maxima merge into one in high temperature region. Thus the phase boundary curve can be drawn in (α, β) plane.

The volume dependency of cluster integrals b_l plays a dominant role in the theory of condensation by cluster integrals. The point of condensation z_{COND} is the point where the gaseous phase and the condensed phase reach equilibrium in the grand canonical en-

---

* Münster[21] gives the criterion of the phase change by the fluctuation in the grand canonical ensemble.

semble. z_{COND} agrees to the irregular point z^* of the limit function $\lim\limits_{N\to\infty}\sum\limits_{l=1}^{\infty} b_l(N) z^l$

$\equiv \lim\left(\sum\limits_{n=0}^{N} n \Omega_n(N) z^n / \sum\limits_{n=0}^{N} \Omega_n(N) z^n\right)$ when $N\to\infty$.

It is doubtful but of no importance whether the smallest real positive singularity z_s of the analytic function $\sum\limits_{l=1}^{\infty} \lim\limits_{N\to\infty} b_l(N) z^l$ agrees with the irregular point of the limit function $\lim\limits_{N\to\infty} \sum\limits_{l=1}^{\infty} b_l(l^r) z^l$ or not. Yang-Lee's theorem does not justify the Mayer's theory.

References

1) L. Onsager, Phys. Rev. **66** (1944), 117.
2) B. Kaufman Phys. Rev. **76** (1949), 1232.
3) Y. Nambu, Prog. Theor. Phys. **5** (1950), 1.
4) C. N. Yang, Phys. Rev. **85** (1952), 80.
5) C. Domb, Proc. Roy. Soc. A **196** (1949), 36, **199** (1949), 199.
6) T. Tanaka, Busseironkenkyu, No. 15, No. 20, (1949). (in Japanese).
7) T. Tanaka, H. Katsumori and S. Toshima, Prog. Theor. Phys. **6** (1951), 17.
8) T. Oguchi, Jour. Phys. Soc. Japan, **5** (1950) 75 ; **6** (1951), 27, 31 ; Busseironkenkyu, No. 22 (1950).
9) A. J. Wakefield, Proc. Cambr. Phil. Soc. **47** (1951), 419, 799.
10) J. E. Mayer, Jour. Chem. Phys. **5** (1937), 67 ; **10** (1942), 629.
     J. E. Mayer and M. G. Mayer, *Statistical Mechanics*, J. Wiley and Sons, (1940).
11) B. Kahn and G. E. Uhlenbeck, Physica, **5** (1938), 399.
12) M. Born and K. Fuchs, Proc. Roy. Soc. A **166** (1938), 391.
13) S. Katsura and H. Fujita, Jour. Chem. Phys. **19** (1951), 799 ; Prog. Theor. Phys. **6** (1951), 498.
14) S. Ono, Busseironkenkyu, No. 38, (1951).
15) H. Fujita and S. Katsura, Sci. Rep. Tôhoku Univ. Ser. I, **35** (1952), 206.
16) K. Ikeda, Busseironkenkyu, No. 57 (1952).
17) S. Ono, Prog. Theor. Phys. **8** (1952), 1. (revised after Ikeda's discussions).
18) J. S. Rushbrooke, Nuovo Cimento **6** (1949), 251.
19) C. N. Yang and T. D. Lee, Phys. Rev. **87** (1952), 404 ; T. D. Lee and C. N. Yang, Phys. Rev. **87** (1952), 410.
20) S. Katsura and H. Fujita, Prog. Theor. Phys. **5** (1950), 997.
21) A. Münster, Proc. Cambr. Phil. Soc. **46** (1950), 319 ; Zeits. f. Naturforschg. **6a** (1951), 139.

Note added in proof Remarks on the Mayer's theory of condensation was discussed by Husimi, Katsura, and Ikeda at the International Conference on Theoretical Physics. (Kyôto, 1953). The content of Ikeda's lecture will be discussed in other place (cf. S. Katsura, Jour. Chem. Phys. **22** (1954)   ).

Small systems have been discussed by A. D. Fokker (Physica **8** (1941), 109.) and P. V. Krishna Iyer (Ann. Math. Stat. **21** (1950), 198). The former is confined to AB-alloy with $N_A = N_B$. The latter has not considered thermodynamic quantities.

In the course of publication of this article, (4×4) etc. were calculated (Lecture on Int. Conf. Theor. Phys. (Kyoto, 1953) ; Busseironkenkyu, No. 70, (Jan. 1954)).

Letters to the Editor

A Variation Principle in Deuteron Problem

Takashi Kikuta

Department of Physics, University of Tokyo

April 9, 1954

In this note it will be shown that, when the binding energy γ^2 and the mixing ratio ε for the D-wave to the S-wave in the deuteron state are given, it is possible to determine the both depths of central and tensor potentials simultaneously by the following variational principle. The excellent computer ILLIAC (University of Illinois, U.S.A.) has been working for the deuteron problem with fruitful results and it seems almost unnecessary to use the methods based on variational principles. However, the procedure given here may be of some interest from a theoretical point of view, being as an extension of Rayleigh-Ritz method to the problems in which two eigenvalues are to be simultaneously determined.

The second order coupled differential equations which arise in the deuteron state are in the following matrix form

$$\boldsymbol{A}(r)\boldsymbol{\psi}(r) = (\mu\boldsymbol{B}(r)+\nu\boldsymbol{C}(r))\boldsymbol{\phi}(r), \qquad (1)$$

where $A_{11}=-d^2/dr^2+\gamma^2$, $A_{22}=-d^2/dr^2+\gamma^2+6/r^2$, $B_{11}=B_{22}=V(r)$, $C_{22}=2W(r)$, $C_{12}=C_{21}=\sqrt{8}W(r)$, $A_{12}=A_{21}=B_{12}=B_{21}=C_{11}=0$, $V(r)$: central well, $W(r)$: tensor well, and $\boldsymbol{\psi}(r)$ has the two components $u(r)$ and $w(r)$. Let the asymptotic form of $\boldsymbol{\psi}(r)$ in the outside region of nuclear potentials be

$$\boldsymbol{\psi}(r) = \begin{cases} u(r) \to \cos\varepsilon\,e^{-\gamma r}, & \gamma > 0, \\ w(r) \to -\sin\varepsilon(3/(\gamma r)^2+3/(\gamma r)+1)e^{-\gamma r}. \end{cases}$$

$$\boldsymbol{\phi}(0)=0. \qquad (2)$$

For the specified binding energy of the deuteron γ^2 and the mixing ratio ε, there are sets of eigenvalues (μ_i, ν_i) in general. We will confine our attention to (μ_0, ν_0) for the ground state. The electric quadrupole moment Q of this state relates ε and is $\varepsilon \doteq -\sqrt{2}\,Q\gamma^2$. For this (μ_0, ν_0) there exists one solution which has the asymptotic form in the outside region of nuclear forces such as

$$\boldsymbol{\phi}(r) = \begin{cases} u(r) \to \sin\varepsilon(\sin h\gamma r + \lambda\cos h\gamma r), \\ w(r) \to \cos\varepsilon(j(\gamma r)+\lambda n(\gamma r)). \end{cases} \quad \boldsymbol{\phi}(0)=0, \qquad (3)$$

$$n(\gamma r) \equiv (1+3/(\gamma r)^2)\cos h\gamma r - (3/(\gamma r))\sin h\gamma r,$$

$$j(\gamma r) \equiv (1+3/(\gamma r)^2)\sin h\gamma r - (3/(\gamma r))\cos h\gamma r.$$

$$\lambda = i\,\tan\delta_2 = \text{real number}.$$

$\boldsymbol{\phi}(r)$ in (2) and (3) will be written as $\boldsymbol{\psi}^{(1)}(r)$ and $\boldsymbol{\psi}^{(2)}(r)$ respectively. Let the approximate functions for $\boldsymbol{\psi}^{(1)}(r)$ and $\boldsymbol{\psi}^{(2)}(r)$ be $\boldsymbol{\phi}^{(1)}(r)$ and $\boldsymbol{\phi}^{(2)}(r)$, which have the same asymptotic form as in (2) and (3) except that, since λ in (3) is unknown, it is replaced by an approximate λ_t. The fundamental formulas in the present method are

$$(\boldsymbol{\phi}^{(1)}, L\boldsymbol{\phi}^{(1)}) - (\varDelta\boldsymbol{\phi}^{(1)}, L\varDelta\boldsymbol{\phi}^{(1)}) = 0, \qquad (4)$$

$$(\boldsymbol{\phi}^{(1)}, L\boldsymbol{\phi}^{(2)}) - (\varDelta\boldsymbol{\phi}^{(1)}, L\varDelta\boldsymbol{\phi}^{(2)}) = 0, \qquad (5a)$$

or

$$(\boldsymbol{\phi}^{(2)}, L\boldsymbol{\phi}^{(1)}) - (\varDelta\boldsymbol{\phi}^{(2)}, L\varDelta\boldsymbol{\phi}^{(1)}) = 0, \qquad (5b)$$

where $L \equiv \boldsymbol{A}-\mu_0\boldsymbol{B}-\nu_0\boldsymbol{C}$, $\varDelta\boldsymbol{\phi}^{(i)} \equiv \boldsymbol{\psi}^{(i)}-\boldsymbol{\phi}^{(i)}$, and ( ) denotes both the inner product for 2-2 matrix and the integration : $\int_0^\infty dr$. These can be derived by noticing that

$$\int_0^\infty \boldsymbol{\phi}^{(i)}\cdot(L\boldsymbol{\phi}^{(j)})\,dr - \int_0^\infty (L\boldsymbol{\phi}^{(i)})\cdot\boldsymbol{\phi}^{(j)}\,dr$$

$$= \frac{d\boldsymbol{\phi}^{(i)}}{dr}\cdot\boldsymbol{\phi}^{(j)} - \boldsymbol{\phi}^{(i)}\cdot\frac{d\boldsymbol{\phi}^{(j)}}{dr}\Big|_0^\infty = 0.$$

Since the terms $(\varDelta\boldsymbol{\phi}^{(i)}, L\varDelta\boldsymbol{\phi}^{(j)})$ in (4) and (5) are order of $\varDelta^2$, we may to the correct order put

$$(\boldsymbol{\phi}^{(1)}, (\boldsymbol{A}-\mu\boldsymbol{B}-\nu\boldsymbol{C})\boldsymbol{\phi}^{(1)}) = 0, \qquad (4')$$

$$(\boldsymbol{\phi}^{(1)}, (\boldsymbol{A}-\mu\boldsymbol{B}-\nu\boldsymbol{C})\boldsymbol{\phi}^{(2)}) = 0, \qquad (5'a)$$

or

$$(\boldsymbol{\phi}^{(2)}, (\boldsymbol{A}-\mu\boldsymbol{B}-\nu\boldsymbol{C})\boldsymbol{\phi}^{(1)}) = 0. \qquad (5'b)$$

Namely, for the trial functions $\boldsymbol{\phi}^{(1)}$ and $\boldsymbol{\phi}^{(2)}$ suitably chosen, *the stationary expressions for μ_0 and ν_0 are given by*

$$\mu = (A_{11}C_{12}-A_{12}C_{11})/(B_{11}C_{12}-B_{12}C_{11}), \qquad (6a)$$

$$A_{ij} \equiv (\boldsymbol{\phi}^{(i)}, \boldsymbol{A}\boldsymbol{\phi}^{(j)}), \text{ etc.}$$

$$\nu = (A_{11}B_{12}-A_{12}B_{11})/(C_{11}B_{12}-C_{12}B_{11}), \qquad (7a)$$

or by similar quantities in which suffixes 1 and 2 are exchanged to each other. In other words, the functions $\boldsymbol{\phi}^{(1)}$ and $\boldsymbol{\phi}^{(2)}$ which satisfy eq. (1) with

the boundary conditions (2) and (3) are just the functions which make the expressions (6) and (7) stationary, and the values of μ and ν computed from such functions give the correct eigenvalues μ_0 and ν_0. (6~7) are an extension of the usual Layleigh-Ritz variational method. If we take the trial functions $\phi^{(1)}$ and $\phi^{(2)}$ containing some adjustable parameters (a_m) and (b_n, λ_t), to make $L_{ij} \equiv (\phi^{(i)}, L\phi^{(j)})$ stationary, which is equivalent to that the expressions (6) and (7) are made stationary, we have to set

$$\partial L_{11}/\partial a_m = 0, \quad \partial L_{ij}/\partial b_n = 0, \quad \partial L_{ij}/\partial \lambda_t = 0.$$

$$(i \neq j, \; i = 1 \text{ or } 2, \; j = 2 \text{ or } 1).$$

The writer is grateful for the kind interest of Prof. T. Yamanouchi.

Renormalization in the Covariant Treatment of Pion-nucleon Scattering

Shin Chiba

*Department of Physics,
Tokyo University of Education*

In the treatment of pion-nucleon scattering in B. S. -formalism,[1] divergencies appear in the solution of integral equations, which are not avoided by the renormalization in the integral kernel.

We show here the method to subtract these divergencies, especially "overlapping divergencies", in a closed form.[2]

In the symmetrical ps-ps theory, we use five integral kernels shown in Fig. 1. For the $T = 3/2$ state, the renormalization is completely achieved in the integral kernel and we do not consider it further. For the $T = 1/2$ state, we have the integral equation for the Green function:
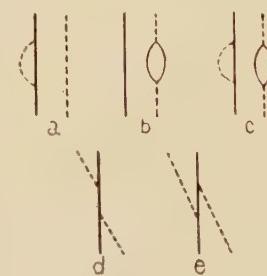
Fig. 1

$$G_{1/2}(p, k; p_0, k_0)$$

$$= \delta(p - p_0)\delta(k - k_0) S_F'(p_0) \Delta_F'(k_0)$$

$$+ S'_F(p) \Delta_F'(k) \cdot \int dp' dk' \delta(P - P')$$

$$\times [3H_1(p, k; p', k') - H_2(p, k; p', k')]$$

$$\times G_{1/2}(p', k'; p_0, k_0), \tag{1}$$

where

$$S_F'(p) = \left[1 - \frac{1}{2\pi} S_{IC}(p)\right]^{-1} S_F(p), \tag{2}$$

$$\Delta_F'(k) = \left[1 + \frac{1}{2\pi i} \Delta_C(k)\right]^{-1} \Delta_F(k), \tag{3}$$

$$H_1(p, k; p', k') = (f^2/2\pi)\gamma_5 S_F(p + k)\gamma_5, \tag{4}$$

$$H_2(p, k; p', k') = (f^2/2\pi)\gamma_5 S_F(p - k')\gamma_5, \tag{5}$$

$$P = p + k, \quad P' = p' + k'. \tag{6}$$

$S_{IC}(p)$ and $\Delta_C(k)$ are finite contributions from second-order self energy parts shown in Fig. 1 (a), (b).

We define $K(p, k; p', k')$ and $\bar{K}(p, k; p', k')$ by

$$K(p, k; p', k') = \delta(p - p')\delta(k - k') - S_F'(p)\Delta_F'(k)$$

$$\times \int dp'' dk'' \delta(P - P'') H_2(p, k; p'', k'')$$

$$\times K(p'', k''; p', k'), \tag{7}$$

$$K(p, k; p', k') S_F'(p') \Delta_F'(k')$$

$$= S_F'(p) \Delta_F'(k) \bar{K}(p, k; p', k'). \tag{8}$$

$K(p, k; p', k')$ and $\bar{K}(p, k; p', k')$ correspond a scattering process as shown in Fig. 2.

Fig. 2

Then, the solution of eq. (1) is easily obtained:

$$G_{1/2}(p, k; p_0, k_0)$$

$$= K(p, k; p_0, k_0) S_F'(p_0) \Delta_F'(k_0)$$

$$+ \delta(P - P_0) S_F'(p) \Delta_F'(k)$$

$$\times \frac{3f^2}{2\pi} v(k; P) [1 - S_F(P) \sum_{II}(P)]^{-1}$$

$$\times S_F(P) w(k_0; P) S_F'(p_0) \Delta_F'(k_0), \tag{9}$$

where

$$v(k; P) = \int dp' dk' \bar{K}(p, k; p', k')\gamma_5, \tag{10}$$

$$w(k; P) = \int dp' dk' \gamma_5 K(p', k'; p, k), \tag{11}$$

$$\sum_{II}(P) = \frac{3f^2}{2\pi} \int d\ell' dk' d\ell'' dk'' \delta(P-P')$$

$$\times \gamma_5 K(\ell', k'; \ell'', k'') S'_F(\ell'') \Delta'_F(k'') \gamma_5. \quad (12)$$

Here, $v(k, P)$ and $w(k, P)$ are vertex contributions obtained by closing one meson line in Fig. 2 and $\sum_{II}(P)$ is a self energy contribution obtained by closing both meson lines in the same figure. Obviously, $v(k; P)$, $w(k; P)$ and $\sum_{II}(P)$ contain "overlapping divergencies". In order to subtract these divergencies, we notice that $v(k; P)$ and $w(k; P)$ satisfy the following integral equations:

$$v(k; P) = \gamma_5 - \int dk' H_2(P-k, k; P-k', k')$$

$$\times S_F'(P-k') \Delta_F'(k') v(k'; P), \quad (13)$$

$$w(k; P) = \gamma_5 - \int dk' w(k'; P) S'_F(P-k')$$

$$\times \Delta_F'(k') H_2(P-k', k'; P-k, k). \quad (14)$$

We replace these equations by the following two sets of equations:

$$v_c(k; P) = \gamma_5 - \int dk' H_2(P-k, k; P-k', k')$$

$$\times S_F'(P-k') \Delta_F'(k') v_c(k'; P),$$

$$+ \int dk' H_2(\ell_0, 0; \ell_0 - k', k')$$

$$\times S_F'(\ell_0 - k') \Delta_F'(k') v_c(k'; \ell_0), \quad (15a)$$

$$v_c(k; \ell_0) = \gamma_5 - \int dk [H_2(\ell_0 - k, k; \ell_0 - k', k')$$

$$- H_2(\ell_0, 0; \ell_0 - k', k')]$$

$$\times S_F'(\ell_0 - k') \Delta_F'(k') v_c(k'; \ell_0), \quad (15b)$$

$$w_c(k; P) = \gamma_5 - \int dk' w_c(k'; P) S_F'(P-k')$$

$$\times \Delta_F'(k') H_2(P-k', k'; P-k, k)$$

$$+ \int dk' w_c(k'; \ell_0) S_F'(\ell_0 - k')$$

$$\times \Delta_F'(k') H_2(\ell_0 - k', k'; \ell_0, 0), \quad (16a)$$

$$w_c(k; \ell_0) = \gamma_5 - \int dk' w_c(k'; \ell_0) S_F'(\ell_0 - k')$$

$$\times \Delta_F'(k') [H_2(\ell_0 - k', k'; \ell_0 - k, k)$$

$$- H_2(\ell_0 - k', k'; \ell_0, 0)], \quad (16b)$$

where

$$\ell_0^2 + m^2 = 0, \qquad \gamma \ell_0 - im = 0. \quad (17)$$

The above defined $v_c(k; P)$ and $w_c(k; P)$ constitute finite vertex contributions. One can easily verify this by solving above equations by iteration.[5]

We can easily find formal solutions of (15a) and (16a):

$$v_c(k; P) = (1-L) v(k; P), \quad (18)$$

$$w_c(k; P) = (1-L) w(k; P), \quad (19)$$

where

$$L = -\frac{f^2}{2\pi} \int dk' S_F(\ell_0 - k') \gamma_5 S_F'(\ell_0 - k')$$

$$\times \Delta_F'(k') v_c(k'; \ell_0)$$

$$= -\frac{f^2}{2\pi} \int dk' w_c(k'; \ell_0) \Delta_F'(k') S_F'(\ell_0 - k') \gamma_5$$

$$\times S_F(\ell_0 - k'). \quad (20)$$

For $\sum_{II}(P)$, we have from (12) and (7)

$$\sum_{II}(P) = \frac{3f^2}{2\pi} \int d\ell' dk' \delta(P-P') [w(k'; P)$$

$$\times S_F'(\ell') \Delta_F'(k') v(k'; P) + \int d\ell'' dk'' \delta(P''-P')$$

$$\times w(k''; P) S_F'(\ell'') \Delta_F'(k'') H_2(\ell'', k''; \ell', k')$$

$$\times S_F'(\ell') \Delta_F'(k') v(k'; P)]. \quad (21)$$

We can obtain "overlapping divergencies"-free self energy contribution, $\sum_{II}'(P)$, replacing $v(k; P)$ and $w(k; P)$ in (21) by $v_c(k; P)$ and $w_c(k; P)$ respectively. By (18) and (19) we obtain

$$\sum_{II}'(P) = (1-L)^2 \sum_{II}(P). \quad (22)$$

(18), (19) and (22) are the relations obtained by Salam.[3] By the usual method we separate diverging parts from $\sum_{II}'(P)$:

$$\sum_{II}'(P) = A_{II} + (\gamma_\mu P_\mu - im) B_{II}$$

$$+ (\gamma_\mu P_\mu - im) S_{IIC}(P). \quad (23)$$

Thus we have a final expression for $G_{1/2}(\ell, k; \ell_0, k_0)$:

$$G_{1/2}(\ell, k; \ell_0, k_0)$$

$$= K(\ell, k; \ell_0, k_0) S_F'(\ell_0) \Delta_F'(k_0)$$

$$+ \delta(P-P_0) S_F'(\ell) \Delta_F'(k)$$

$$\times \frac{3f^2}{2\pi} v_c(k; P) \left[1 - \frac{1}{2\pi} S_{IIC}(P) \right]^{-1}$$

$$\times S_F(P) w_c(k_0; P) S_F'(\ell_0) \Delta_F'(k_0). \quad (24)$$

No divergencies appear in this final formula.[4]

1) S. Fubini, Nuovo Cimento 10 (1953), 851.
2) Recently, the author was shown a letter sent by M.M. Lévy to S. Tomonaga. In this letter, Lévy describes a method of subtracting "overlapping divergencies" in nearly the same manner as the author. Also, D. Ito and H. Tanaka treat same problem in some different way.

3) A. Salam, Phys. Rev. **82** (1951), 217.

4) We have finished the calculation of phase-shifts in pion-nucleon scattering, applying the above method of renormalization to the Tamm-Dancoff approximation. The details will soon appear in this journal.

Further Comments on the Proposed Mass Spectrum

Sadae Yoshikawa and Tomimaro Hasebe

Mito Branch of Hitachi Works, Hitachi, Ltd. Katsuta-cho, Ibaragi-ken

April 17, 1954

Our previously proposed mass spectrum[1] can be written also as follows .

$$\mu_n = (\hbar c/e^2)\,\omega^n.$$

Here notations are the same as before and $\omega = (\hbar c/e^2)^{1/27.5}$. Since the precise value of $\hbar c/e^2$ is known as 137.036, the results calculated with this value are listed in Table 1,[2–8] where the mass values smaller than that of the deuteron are shown.

As remarked by Fang,[9] the above calculation has the advantage that the nucleon mass will be effected appreciatively by taking a more realistic value of 137.036 instead of 137, which may be attributed to some "meson effect". In fact, if what is meant by the predicted value of $\mu_{14\,1/2} = 1834.55m_e$ being smaller than the observed value of the proton mass only by about $1.5m_e$ should happen to be related to the well-known problem that the simplest formula concerning the electro-magnetic characteristic of nucleons gives the inverse sign against observation, the numerical consideration would be of great significance.

The subsequent accumulation of important experimental data has necessitated the revision of the table of comparison between the calculated and observed values as shown in Table 1. It may deserve to be mentioned especially that Thompson et al.[5] made the V_1^0-particle decay of $V_1^0 \to p + \pi^- + \sim 37$Mev definitely conclusive by verifying that the Q curve plot of the V_1^0 disintegration data clearly forms itself into the expected ellipse in the so-called (a, p_T) plane. Hereby our mass spectrum has added one

certainly reliable evidence for the correspondence of n to spin; viz, as for the V_1^0-particle mass, the corresponding value is given by $\mu_{15\,1/2} = 2194m_e$ from our mass spectrum whereas $\mu_{16.0} = 2192m_e$ from Nambu's.[1,10] However, at present things do not always favor us. For instance, a decay of the type suggested by Peters et al.,[3] namely $V^\pm \to \pi^\pm + (\text{neutron}) + \sim 135$ Mev, somehow appears to contradict the above correspondence (see Table 1), though many evidences for this type are of course necessary before the adequacy of the assumption of this type can be considered.

Leprince-Ringuet and Rossi[11] have emphasized that one must maturely consider whether several different kinds of K mesons actually exist. However, it must be considered also that, from the Lorentz-invariance, the τ and V_4^0[12] particles are theoretically impossible to be the counterpart of each other, provided the V_4^0-particle decay is surely unrelated to the photon emission.[13] Further, heterodox as it may sound, we wonder if two of the six K mesons observed by the Paris Group,[14] e.g., the K_{ρ_2} and K_{ρ_3} particles whose masses were measured to be $(850\pm150)m_e$ and $(1065\pm160)m_e$ respectively by scattering range, could in truth be of the same mass of $(940\pm40)m_e$, for these values can be presumed also to correspond to the values of $\mu_{10\,1/2} = 897m_e$ and $\mu_{11\,1/2} = 1073m_e$ respectively. En passant, it seems natural to consider that, as marked previously, the charged ζ meson differs in type from the neutral ζ meson because of the excessive disparity in their lifetime, even if the former particle decay can not be assumed as $\zeta^\pm \to \mu^\pm + 2\nu$. Two events of charged ζ mesons of the similar type were recently reported also by Seeman et al.[15]

We have many experimental facts comparable with the case of the positive n as shown in Table 1, while few ones comparable with the case of the negative n. However, it may be apparently overhasty to say that we have no evidences in need of careful thought at all for the existence of particles less massive than the μ meson, except such light particles as leptons and photons. For example, formerly Brode[16] found two different kinds of light mesons whose respective masses were definitely near upon the values of 100 and $70m_e$. On the other hand, Schein et al.[17] have recently reported the production of high energy "electron pairs" of mass less than $10m_e$ by operating the 450Mev Chicago cyclotron. And they confess in their report, "It is not entirely impossible that these particles are composed of some new particles of mass less than $10m_e$."

Table 1. Comparison between the calculated and observed values, when $18 \geqq n \geqq 0$, $n = -27\frac{1}{2}$, and $n = -\infty$.

Calculation		Experimental Facts					
n	Mass (m_e)	Particles	Assumed mode of decay	Q (Mev)	Spin $(\hbar)$	Mass (m_e)	Reference No.
18.0	3431.55						
$17\frac{1}{2}$	3137.90						
17.0	2869.38						
$16\frac{1}{2}$	2623.85	Y^-	$Y^- \rightarrow V_1^0 + \pi^-$	~ 60	$(\frac{1}{2})$	~ 2600	2
16.0	2399.30	$V^\pm$	$V^\pm \rightarrow (n)^* + \pi^\pm$	135 ± 35	$(\frac{1}{2})^*$	2330 ± 30	3
$15\frac{1}{2}$	2193.98	V^+ V_1^0	$V^+ \rightarrow p + \pi^0$ $V_1^0 \rightarrow p + \pi^-$	~ 40 37 ± 2	$(\frac{1}{2})$	~ 2190 2184 ± 4	4 3.5
15.0	2006.23						
$14\frac{1}{2}$	1834.55	p n			$\frac{1}{2}$	1836.1 1838.6	
14.0	1677.56						
$13\frac{1}{2}$	1534.00						
13.0	1402.73	V_3^0	$V_3^0 \rightarrow (\tau^-) + \pi^+$	~ 60	(0.1)	~ 1370	6
$12\frac{1}{2}$	1282.69	The corresponding particles may be identified with some of those which are called K mesons such as χ, $\varkappa$, τ mesons, etc. One example of $\varkappa$ mesons was shown in such a fashion as listed next.					
12.0	1172.93						
$11\frac{1}{2}$	1072.55	$\varkappa^\pm$	$\varkappa^\pm \rightarrow \mu^\pm + 2(\nu)$	Continuous	$(\frac{1}{2})$	1080 ± 100	7
11.0	980.77	$\tau^\pm$	$\tau^\pm \rightarrow \pi^\pm + \pi^+ + \pi^-$	76 ± 3	(0.1)	977 ± 5	7
$10\frac{1}{2}$	896.84						
10.0	820.10	V_2^0	$V_2^0 \rightarrow \pi^+ + \pi^-$	132 ± 17	(0.1)	810 ± 35	3
$9\frac{1}{2}$	749.92						
9.0	685.74						
$8\frac{1}{2}$	627.06						
8.0	573.40	ζ^0 $\zeta^\pm$	$\zeta^0 \rightarrow \pi^+ + \pi^-$ $\zeta^\pm \rightarrow \pi^\pm + \pi^0$	~ 2 ~ 1	(0.1)	556±4 535±35	7
$7\frac{1}{2}$	524.33						
7.0	479.46						
$6\frac{1}{2}$	438.43						
6.0	400.92						
$5\frac{1}{2}$	366.61						

* What is known as to the neutral secondary of the $V^\pm$ is the fact that its mass must be larger than $1570 m_e (< \mu_{14.0} = 1677.56 m_e)$. See ref. 3.

5.0	335.24						
4½	306.55						
4.0	280.32	$\pi^{\pm}$ π^{0}	$\pi^{\pm}\to\mu^{\pm}+\nu$ $\pi^{0}\to2\gamma$	~34	0	~276 ~265	8
3½	256.33						
3.0	234 39						
2½	214.33	$\mu^{\pm}$	$\mu^{\pm}\to e^{\pm}+2\nu$	Continuous	½	~210	8
2.0	195.99						
1½	179.22						
1.0	163.88						
½	149.86						
0	137.036						
−27½	1.000	$e^{\pm}$			½	1	
−∞	0.000	ν			½	~0	
		γ			1	0	

Now we would like to express our gratitude to Profs. M. Inoki and T. Yasaki of Yamanashi University for informing us, prior to publication, of their hopeful effort to search for such light mesons and other unknown cosmic ray particles, now in progress at their laboratory. We are also indebted to Dr. P. H. Fang of the Catholic University of America at Washington for his kind advice and interest in our work.

1) S. Yoshikawa and T. Hasebe, Prog. Theor. Phys. 10 (1953), 359.

2) C. D. Anderson, E. W. Cowan, R. B. Leighton, and van Lint, Bull. Am. Phys. Soc. 28 No. 5 (1953), 17.

3) D. Lal, Yash Pal and B. Peters, Phys. Rev. 92 (1953), 438.

4) C. M. York, R. B. Leighton, and E. K. Bjornerud, Phys. Rev. 90 (1953), 167.

5) R. W. Thompson, A. V. Buskirk, L. R. Etter, C. J. Karzmark, and R. H. Rediker, Phys. Rev. 90 (1953), 329.

6) R. B. Leighton, S. D. Wanlass, and C. D. Anderson, Phys. Rev. 89 (1953), 148.

7) C. F. Powell, Report of the International Physics Conference (Copenhagen), June 3–17 (1952), p. 27.

8) F. M. Smith, W. H. Barkas, A. S. Bishop, H. Bradner, and E. Gardner, Bull. Am. Phys. Soc. 24, No. 8 (1949), 9; R. E. Marshak, Rev. Mod. Phys. 23 (1951), 131.

9) P. H. Fang, Private Communication.

10) Y. Nambu, Prog. Theor. Phys. 7 (1952), 595.

11) L. Leprince-Ringuet and B. Rossi, Phys. Rev. 92 (1953), 722.

12) The symbol V_4^0 is used to designate the V^0 particle of which the decay scheme is assumed as $V_4^0\to\pi^{+}+\pi^{-}+\sim210\text{Mev}$. See Reference 5.

13) H. Fukuda, Discussion Remark at the International Physics Conference (Kyoto), September 18–23 (1953).

14) J. Crussard, L. Leprince-Ringuet, D. Morellet, A. Orkin-Lecourtois, and J. Trembley, Phys. Rev. 90 (1953), 1127

15) N. Seeman, D. T. King, and M. M. Shapiro, Phys. Rev. 91 (1953), 446.

16) R. B. Brode, Rev. Mod. Phys. 21 (1949), 37.

17) M. Schein, J. Fainberg, D. M. Haskin, and R. G. Glasser, Phys. Rev. 91 (1953), 973.

On the x-μ Decay-in-flight in the Photographic Emulsion exposed at High Altitude

Kiyosi Niu

Cosmic Ray Laboratory, Physical Institute, Nagoya University

April 24, 1954

During the last few years a number of examples[1]−[4] have been reported of heavy mesons which are usually referred as x or χ particles. Each of those unstable heavy particles decays near the end of its range into single lighter charged meson.

Recently, in our laboratory, a particle, which can not be identified as a proton or a π-meson and decays in flight into a lighter meson, has been found among star particles in a photographic plate exposed at high altitude. The plate in which this particle was found was one of those which were exposed and processed by Dr. V. Perterson*. They are $3'' \times 3''$, 400μ, Ilford G–5 emulsions, and flown by means of free balloons which reached a minimum pressure of 1.2 m.m. Hg, at Bismark, North Dakota (magnetic latitude 56° N) on August 3, 1951.

The event in question is a simple star composed of two thin tracks, one grey track and one black track. Fig. 1. shows the projections of the star on the planes parallel and normal to the emulsion-glass
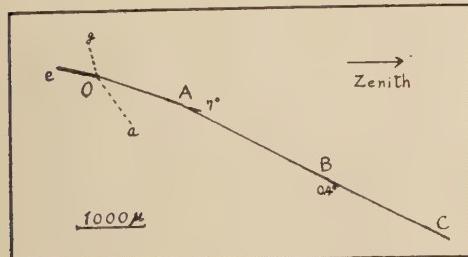

Fig. 1. (A) The projection of the star on the plane parallel to the emulsion-glass plane.

Oa; Thin track emerging to the surface, grain density $25/100\mu$.

Og; Thin track emerging to the glass, grain density $34/100\mu$.

Oe; Black track stopping in the emulsion, about 500μ long.

OABC; Grey track.

Fig. 1. (B) The projection of the star on the plane normal to the emulsion-glass plane.
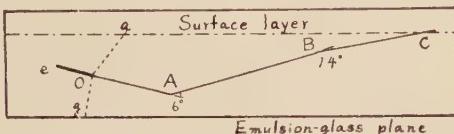
Table I. The corrected values of each part of the grey track.

	OA	AB	BC
Length μ	1217	2326	1998
Grain density grains/100μ	72.1	76.1	43.7
Mean scattering angle $°/(100\mu)^{1/2}$	0.195	0.195	0.336

plane. Two thin tracks have very large dip angles as are shown in Fig. 1. (B), and the primary particle of this event cannot be identified. The grey track shows a sudden change of direction about 7° at the point A and furthermore abrupt decrease of its grain density at the point B, and then enter into the surface layer** at the point C. However, the direction change at the point B is very small, about 0.4° parallel and about 1.4° normal to the emulsion-glass plane. The length, grain density and mean scattering angle of each part of this track are shown in Table I, after correction for distortion which was found in the region of the plate near the present event. Relations between the grain density and the scattering angle of the grey track and those of the other tracks in this plate are shown in Fig. 2. From this figure one can see that the particle which produced the grey track has the mass near the proton mass in the part OB and near the π-meson mass in the part BC respectively.

In the first place one might assume that the particle in OB would be a proton and that in BC a π-meson. Then their kinetic energies should be about 70 Mev and 40 Mev respectively. This makes the interpretation of the whole event quite difficult under any assumption on the directions of motion of these particles, because a proton whose kinetic energy is 70 Mev can neither produce a meson nor a star which has two thin tracks. Next let us assume that the particle in OB is one of unstable heavy

Fig. 2. Relation between the grain density and the scattering angle.

; The grey track.
; The tracks related to stars.
; Single-tracks in this plate.
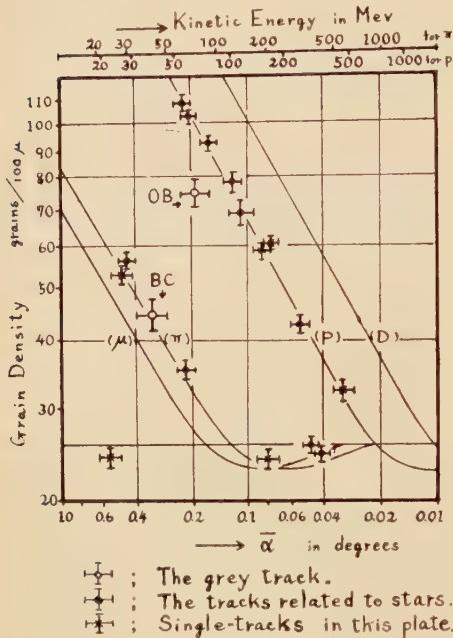
mesons, its mass (1160 ± 170me) being close to the proton mass but smaller than it. Then the event can be interpreted as follows. An unstable heavy particle was produced in the star and decayed in flight into lighter meson (264 ± 55me) at the point B. Then the kinetic energy of the light meson in the rest system of the parent particle is about 7 Mev. This value is inconsistent with the current knowledge about the unstable heavy meson χ[5] which is considered to decay into one neutral particle and one charged π-meson with unique kinetic energy of about 110Mev. If one assumes two particle decay mode, therefore one should introduce another type of heavy mesons. Alternatively, however, if we assume three particle decay mode, the present event can be interpreted as a χ-μ decay. Because the observed momentum value of the secondary particles is not inconsistent with the expected one which ranges continuously from zero to the maximum value of about 220 Mev/c[3], provided that three particle decay mode (one charged μ-meson and two neutral particles) is assumed on the χ particle. The probability that the parent particle decays in the course OB is not so small, on account of the time of flight of 0.26×10^{-10} seconds.

Usually, the determination of such an event depends on the abrupt change in grain density at a given point and the direction change at the same point. In the present event the crucial evidence is based only upon the abrupt change in the grain density at B, while the direction change at B is very small. Alternatively, therefore we examined whether this event will occur due to a spurious distortion such as "wetting effect"[6] or not. In the presnt case the wetting effect will not completely be neglected. However, the order of the abrupt decrease in grain density at B seems quite larger than that expected from the wetting effect, (Fig. 3), and it is presumed that the probability of observing this type of
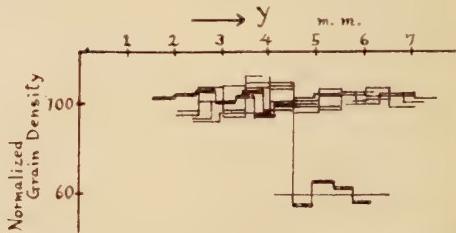

Fig. 3. The result of the examination of the possible local variation of the grain density related to the y-coordinate in Fig. 1. The grain density of each track is normalized by its mean value.
—; The grey track.
—; Neighbouring tracks.

phenomena due to the wetting effect is much smaller than that of observing this type of phenomena produced by the χ-μ decay[6]. After all, it seems to us most reasonable to explain this event as a χ-μ decay-in-flight.

The author wishes to express his appreciation to Dr. V. Peterson for his kindness of providing these valuable plates to our laboratory, and to the "Office of Naval Research, U.S. Navy" for financial support to the balloon flight program. He is also indebted to Prof. Y. Sekido, Prof. I. Miura, Mr. Y. Tuzuki and Mr. J. Nishimura for many helpful discussions.

1) C. Oceallaigh, Phil. Mag. **42** (1951), 1032.
2) Levi-setti and Tomasini, Nuovo Cimento **IX** (1952), 1224.
3) J. Crussard, L. Leprince-Ringuet, D. Morellet, A. Orkin-Lecourtois and J. Trembley, Phys. Rev. **93** (1953), 1127.
4) N. Isachsen, V. Vangen and S. O. Sörensen, Phil. Mag. **44** (1953), 224.
5) M. G. K. Menon, International Conference on Cosmic Radiations, Bagnères de Bigorre, France (1953).
6) K. Niu, (to be published on the Journal of the Physical Society of Japan).

Covariant Subtraction of "Overlapping Divergences" appearing in the Pion-nucleon Scattering

Daisuke Ito and Hiroshi Tanaka

Physics Institute,
Tokyo University of Education

April 27, 1954

In this brief note, we shall discuss a possible way to subtract the so called "overlapping divergences" appearing in the solution of Bethe-Salpeter equation[1] for the Pion-Nucleon scattering in accordance with Dyson-Salam's prescriptions,[2] and to construct the finite part of scattering matrix elements in a closed form.

The B-S eq. for the pion-nucleon scattering under consideration may be written as

$$\psi = \psi_0 + (K_1 + K_2)\psi, \qquad (1)$$

where K_1, K_2 are integral kernels of the form illustrated in Fig. 1. The formal solution of this equation may be expressed, as shown below, with the aid of the resolvent $1/(1-K_2)$ of kernel K_2 in the following form

$$\psi = \frac{1}{1-K_2}\left[1 + \frac{1}{F}H^*S_F\frac{1}{1-\Sigma \cdot S_F}H\frac{1}{1-K_2}\right]\psi_0, \quad (2)$$

where $1/F$ is a two body propagator of the meson-nucleon system, and H^*, H are simple vertices of creating and annihilating a pion, respectively. (In the Ps-Ps theory they are $ig\gamma_5$). Σ is a self energy operator of the following form

$$\Sigma = H\frac{1}{1-K_2}\frac{1}{F}H^* \equiv H\frac{1}{F-I_2}H^*, \qquad (3)$$

where $I_2 \equiv FK_2$ is a well known interaction kernel.

Eq. (2) can easily be verified, if we take into account the structure of the kernel $K_1 \equiv 1/F \cdot H^*S_FH$. Inserting this into (1), we have[2]

$$\psi = \frac{1}{1-K_2}\psi_0 + \frac{1}{1-K_2}\frac{1}{F}H^*S_FH\psi. \qquad (4)$$

Operating H from the left and solving with respect to $H\psi$, we have

$$H\psi = \frac{1}{1-H\frac{1}{1-K_2}\frac{1}{F}H^*S_F}H\frac{1}{1-K_2}\psi_0.$$

Inserting this into (4), eq. (1) follows.

The importance of this result is that the solution of eq. (1) of the π-N scattering can be obtained in a closed form, if we are able to find out the resolvent $T = 1/(1-K_2)$ by solving a more simple divergence-free equation $(1-K_2)T = 1$ in a closed form.

However, we shall encounter divergences in the course of constructing ψ from the resolvent of K_2; i.e., a self energy operator (3) and vertex operators

$$\Gamma = H\frac{1}{1-K_2}, \qquad \Gamma^* = \frac{1}{1-K_2}\frac{1}{F}H^* \qquad (5)$$

appearing in (2) contain the so called "overlapping divergences" which should be removed by Dyson-Salam's prescriptions.

In order to remove the divergence from the vertex-operator Γ, we divide the kernel K_2 of the integral equation for $\Gamma(p; k)$ (see Fig. 2)

$$\Gamma(p; k) = H + \int \Gamma(p; k')dk'K_2(p; k'k) \quad (6)$$

into two parts, $K_2(p; k'k) = K_2(p_0; k'k) + [K_2(p; k'k) - K_2(p_0; k'k)]$, and denote the first and the second term of this separation as $K_0(k'k)$ and $K'(p; k'k)$, respectively, where p_0 is a free nucleon momentum $p_0{}^2 + m^2 = 0$.

Then the formal solution of $\Gamma = H + \Gamma(K_0 + K')$ is given

$$\Gamma = H\frac{1}{1-K_0} + H\frac{1}{1-K_0}K'\frac{1}{1-K_2} \equiv \Gamma_0 + \Gamma_0K'\frac{1}{1-K_2}. \qquad (7)$$

Dividing the kernel K_0 of the integral equation satisfied by $\Gamma_0 \equiv H\frac{1}{1-K_0}$ into two parts, $K_0(k'k) = K_0(k'0) + [K_0(k'k) - K_0(k'0)] \equiv \overleftarrow{K_0} + K_0'$, we have the formal solution for Γ_0

$$\Gamma_0 = [H + (\Gamma_0\overleftarrow{K_0})]\frac{1}{1-K_0'}, \qquad (8)$$

where $(\Gamma_0\overleftarrow{K_0}) \equiv \int\Gamma_0(k)dkK_0(k0)$ contains no k-variable. To calculate this unknown operator $(\Gamma_0\overleftarrow{K_0})$, we operate $\overleftarrow{K_0}$ from the right and we get

$$(\Gamma_0\overleftarrow{K_0}) = [H + (\Gamma_0\overleftarrow{K_0})]\left(\frac{1}{1-K_0'}\overleftarrow{K_0}\right).$$

As readily seen,

$$L \equiv \left(\frac{1}{1-K_0'}\overleftarrow{K_0}\right) = \int K_2(p_0; k0)dk$$

$$+ \iint K_0'(kk')K_2(p_0; k'0)dkdk' + \cdots \quad (9)$$

is a divergent constant (pure number), and also identical with Salam's L.

Therefore $(\Gamma_0 K_0') = \dfrac{L}{1-L} H$ holds, and inserting this into the above expression of Γ_0, we have

$$\Gamma_0 = \frac{1}{1-L} H \frac{1}{1-K_0'} \equiv \frac{1}{1-L} \Gamma_{0c}, \qquad (10)$$

where $\Gamma_{0c} \equiv H \dfrac{1}{1-K_0'}$ is free from divergences, as eradily proved. Introducing this back into (7) we obtain finally

$$\Gamma = \frac{1}{1-L} \left(\Gamma_{0c} + \Gamma_{0c} K' \frac{1}{1-K_2'} \right) \equiv \frac{1}{1-L} \Gamma_c, \quad (11)$$

where the finiteness of Γ_c is evident, for Γ_{0c} is finite and $K' \equiv K_2(p; k'k) - K_2(p_0; k'k)$ eliminates the source of all overlapping divergences. Eq. (11) shows that the finite part thus obtained is constructed from the resolvents $\dfrac{1}{1-K_0'}$ and $\dfrac{1}{1-K_2'}$ in a closed form.

In order to separate the divergences in the self energy operator (3), we rewrite $F-I$ as $F-I = (F_0 - I_0) + (F' - I')$, where $F_0 \equiv F(p_0;\ )$ and $I_0 = I_2(p_0;\ )$. Then

$$\Sigma = H \frac{1}{F_0 - I_0} H^* + H \frac{1}{F_0 - I_0} (F' - I') \frac{1}{F_0 - I_0} H^*$$

$$+ H \frac{1}{F_0 - I_0} (F' - I') \frac{1}{F - I_2} (F' - I') \frac{1}{F_0 - I_0} H^*. \tag{12}$$

Owing to the fact $F' = I' = 0$ for $p = p_0$, $\Sigma(p_0)$ is equal to $H \dfrac{1}{F_0 - I_0} H^* \equiv \Sigma_0$ which should be compensated by the counter term of nucleon mass renormalization. By making use of the following relations and the associated ones

$$\frac{1}{F_0 - I_0} = \frac{1}{1 - K_0} F_0,$$

$$H \frac{1}{F_0 - I_0} \equiv \Gamma_0 \frac{1}{F_0} = \frac{1}{1-L} \Gamma_{0c} \frac{1}{F_0},$$

and

$$\frac{1}{F_0 - I_0} H^* \equiv \Gamma_0^* = \frac{1}{1-L} \Gamma_{0c}^*,$$

we have

$$\Sigma = \Sigma_0 + \frac{1}{(1-L)^2} \Big[\Gamma_{0c} \frac{1}{F_0} (F' - I') \Gamma_{0c}^*$$

$$+ \Gamma_{0c} \frac{1}{F_0} (F' - I') \frac{1}{F - I_2} (F' - I') \Gamma_{0c}^* \Big]$$

$$\equiv \Sigma_0 + \frac{1}{(1-L)^2} [\Sigma_1^* + \Sigma_2^*],$$

where $\Sigma_1^*(p)$ still contains the "true divergence" $\left(\dfrac{\partial \Sigma_1^*}{\partial p} \right)_{p=p_0}$. Separating out this divergence we finally obtain the finite part of the self energy operator, which is also constructed from the resolvent operators $\dfrac{1}{F_0 - I_0}$ and $\dfrac{1}{F - I_2}$.
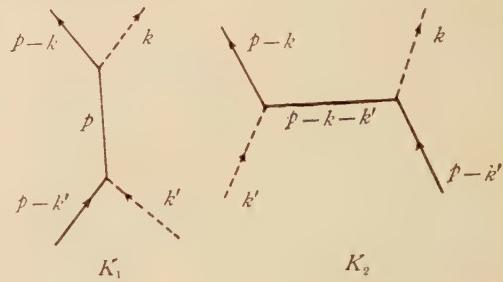

Fig. 1.
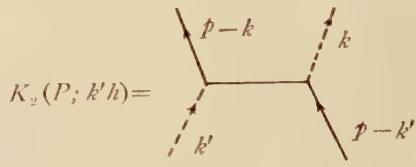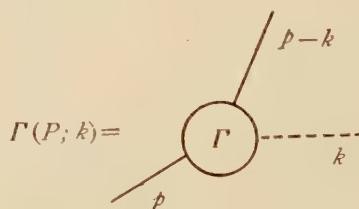
Fig. 2.

* The content of this letter is read at the meeting of Phys. Soc. Japan held on April 2, 1954.

1) E. E. Salpeter and H. A. Bethe, Phys. Rev. **84** (1951), 1232.
2) A. Salam, Phys. Rev. **82** (1951), 217.
3) This result is essentially the same as that derived by S. Fubini, Nuov. Cim. **10** (1953) 685.

Additional remark

After the completon of this work, we were informed of Professor Lévy's recent work on the same subject, through his private communication to Professor Tomonaga. In his work which will soon appear in Phys. Rev., the result of vertex part renormalization is identical with ours, but his method of determining the finite part of vertex operator is somewhat different from ours. His integral equation for determining

the finite part seems to contain a difference of divergent integrals of unknown operators. His treatment for self energy renormalization may be rather different from our method.

His divergent quantity $\varLambda^{(\alpha)}(f_0, f_0)$ may be expressed in our notation as $\left(\dfrac{1}{1-K_0} \overset{\leftarrow}{K_0}\right)$ and it is readily shown, through direct calculation, that his $1 + \varLambda^{(\alpha)} \cdot (f_0, f_0)$ is identical with our (Salam's) $\dfrac{1}{1-L}$.

Further Considerations Regarding Liquid Helium II at High Pressures

Ziro Mikura

The Research Institute for Iron, Steel and Other Metals, Tohoku University,

May 10, 1954

In order to discuss the properties of He³-He⁴ mixtures and those of pure He⁴ at high pressures, a "modified Bose-Einstein liquid" theory was proposed in previous papers[1],[2], in which the energy of excitation, $\varDelta$, and the mass factor, ν, were determined by the relations

$$\varDelta = \varDelta_0 \frac{n_4/\nu}{n_4^0/\nu_0}\left(\frac{M_4}{M_4^0}\right)^{1/3}, \tag{1}$$

$$\nu \equiv M_4/m_4 = \nu_0 + 0.15(f - f_0), \tag{2}$$

where n_4 is the number density of M_4 atoms, M_4 the mass of a He⁴ particle, m_4 the mass of a He⁴ atom, and p the pressure in atm. All the quantities with (0) refer to some standard state. When the state corresponding to the normal lambda-point ($f_0 = 0.05$, $n_4^0 = 0.1462/m_4$)* was chosen as this reference state, we found, as the most appropriate values, $\nu_0 = 8.8$ and $\varDelta_0 = 8.609°K$. For the particular case of pure He⁴ at high pressures, eqs. (1) and (2) then lead to

$$\varDelta = 8.609 \frac{\rho}{0.1462}\left(\frac{\nu}{8.8}\right)^{2/3}, \tag{3}$$

$$\nu = 8.8 + 0.15(f - 0.05). \tag{4}$$

* 0.1462 g/cc is the density of liquid He⁴ at the normal lambda-point.

In this note, the results of calculation of specific heat and thermal Rayleigh disk torque at high pressures are presented.

The specific heat at high pressures has already been calculated in a previous paper[2], but comparison with experiment was not done as I had overlooked the experimental work by W. H. Keesom and Miss Keesom[3] on the specific heat at 19 and 25 atm. In Fig. 1, the specific heat calculated by employing for $\varDelta$ and ν the values from (3) and (4) is compared with this experiment. Agreement is almost satisfactory.


Fig. 1. Specific heats of liquid He⁴ at high pressures. Experimental data (Keesom and Keesom[3]) are corrected so as to be in the agreed temperature scale.

Applying the same model, the effect of pressure on the thermal Rayleigh disk torque is also calculated. As Pellam and Hanson[4] has shown, the torque is given by

$$\tau = \frac{4}{3}a^3 \rho_n \left(\frac{\rho_n}{\rho_s}\right)\frac{\langle H^2\rangle_{Av}}{(\rho ST)^2},$$

where a is the diameter of the disk, H the heat current, and the other quantities are of the usual meanings. The torque at high pressures relative to that under the saturated vapour pressure is then

$$\frac{\tau}{\tau_0} = \frac{\rho_0}{\rho}\frac{x}{x_0}\frac{1 - x_0}{1 - x}\left(\frac{S_0}{S}\right)^2, \tag{5}$$

where $x \equiv \rho_n/\rho$ and the suffix (0) refers to the liquid under the saturated vapour pressure. For ρ_0 and ρ in Eq. (5) the empirical values are used, and for x_0, x, S_0, and S the values obtainable from

$$x = \frac{\nu m_4}{\rho} \left(\frac{2\pi \nu m_4 k T}{h^2} \right)^{3/2} \exp(-\Delta/T),$$

$$S = \frac{k}{\rho} \left(\frac{2\pi \nu m_4 k T}{h^2} \right)^{3/2} \left(\frac{\Delta}{T} + \frac{5}{2} \right) \exp(-\Delta/T)$$

are used. The values Δ and ν are derived from eqs. (3) and (4). They differ but little from the values listed in Table II of reference 2, which were obtained
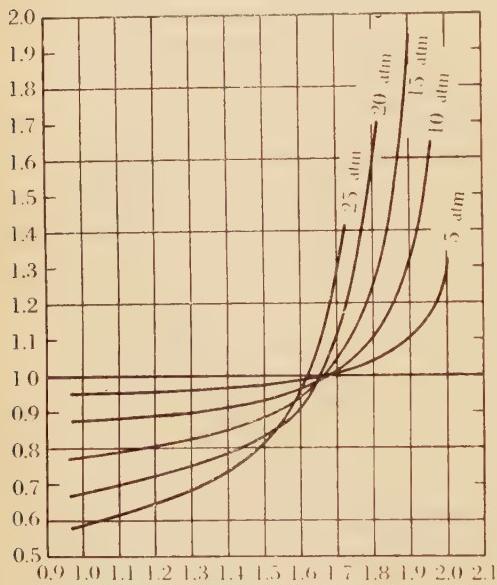
Fig. 2. Thermal Rayleigh disk torque at high pressure, relative to that under saturated vapour pressure, plotted as a function of temperature.

from eq. (3) without using eq. (4) but by requiring the calculated transition temperature to be in accord with experiment. The results of calculation are shown in Fig. 2.

1) Z. Mikura, Prog. Theor. Phys. **11** (1954), 25.
2) Z. Mikura, Prog. Theor. Phys. **11** (1954), 207.
3) W. H. Keesom and A. P. Keesom, Physica **2** (1935), 557.
4) J. R. Pellam and W. B. Hanson, Phys. Rev. **85** (1952), 216.

On the Thermal Rayleigh Disk in Liquid Helium Containing He[3]

Ziro Mikura

*The Research Institute for Iron,
Steel and Other Metals,
Tohoku University,*

The effect of He[3] on the thermal Rayleigh disk torque in liquid helium was first discussed by Koide and Usui[1] on the basis of the de Boer-Gorter-Taconis model of a He[3]-He[4] mixture. They assumed, as they did in discussing the second sound velocity, that He[3] is stationary with the center of gravity system. They appear to have concluded that the marked decrease in torque resulting from the addition of He[3] is intimately connected with the above mentioned special assumption concerning the motion of He[3] in second sound propagation. Recently the author proposed a "modified Bose-Einstein liquid" theory[2] of a He[3]-He[4] mixture, which makes it possible to explain the increase in second sound velocity arising from the dilution without making use of Koide and Usui's assumption but by employing a more plausible assumption: He[3] moves with the same velocity as the normal fluid. In the present note, the effect of the dilution on the thermal Rayleigh disk torque is discussed from the view point based on the same ground.

The torque τ in a mixture is easily found to be

$$\tau = \frac{4}{3} \frac{a^3}{\rho} \frac{\rho_{n4} + \rho_3}{\rho_s T'^2} \frac{\langle \dot{H}^2 \rangle_{Av}}{\{(1-x_3) \widetilde{S}_4 + x_3 S_3\}^2}, \quad (1)$$

where

$$\widetilde{S}_4 = \frac{k}{\nu m_4} \left(\frac{\Delta}{T} + \frac{5}{2} \right) x, \quad (2)$$

$$S_3 = \frac{k}{m_3} \left\{ -\ln C_3' + \ln 2 V_3^0 \left(\frac{2\pi m_3 k T}{h^2} \right)^{3/2} + \frac{5}{2} \right\},$$

$$C_3' = N_3 V_3^0 / (N_3 V_3^0 + N_4 V_4^0), \quad (3)$$

and a is the diameter of the disk, $\dot{H}$ the heat current. For the meanings of the other notations, readers are referred to the previous paper[2]. The expression (3) for S_3 is based on the assumption that He[3] in the mixture liquid can be treated as a nondegenerate Fermi-Dirac gas. It is to be noticed that this assumption can be successfully applied to the problem on vapour pressure and second sound velocity in the

mixture[2]). From eq. (1), the relative value of the torque in a mixture to that for pure He^4 under the same conditions is easily found to be

$$\frac{\tau}{\tau_0} = \frac{1+(I_3^{\sim 0}/I_4^{\sim 0}-1)\xi_3}{1-\xi_3/4}\frac{1-x_0}{x_0}\frac{1-(1-x_3)(1-x)}{(1-x_3)(1-x)}$$

$$\left(\frac{\widetilde{S_4^0}}{\widetilde{S_4}}\right)^2\frac{1}{\{1-x_3(1-\widetilde{S_3}/S_4)\}^2}, \tag{4}$$

where ξ_3 is $N_3/(N_3+N_4)$. Calculations were performed by employing for ν and $\varDelta$ in eq. (2) the values obtainable from the theory in reference 2, for $\widetilde{S_4^0}$ the observed value $[S_4^0]_{obs.}$, for x_0 the semi-empirical value $[x_0]_{obs.}$ derivable from the observed values of S_4^0, C_p^0, and v_{II}^0, and finally for x and $\widetilde{S_4}$ those calculated from

$$x = [x_0]_{obs.} \times (x/x_0)_{theor.},$$

$$\widetilde{S_4} = [S_4^0]_{obs.} \times (\widetilde{S_4}/\widetilde{S_4^0})_{theor.}. \tag{5}$$

Fig. 1 shows the results of this calculation compared with those obtained from Koide and Usui's theory. The values of x employed in the latter calculation were corrected similarly by using eq. (5). Both theories result in qualitatively the same conclusion that the torque must decrease markedly by the dilu-
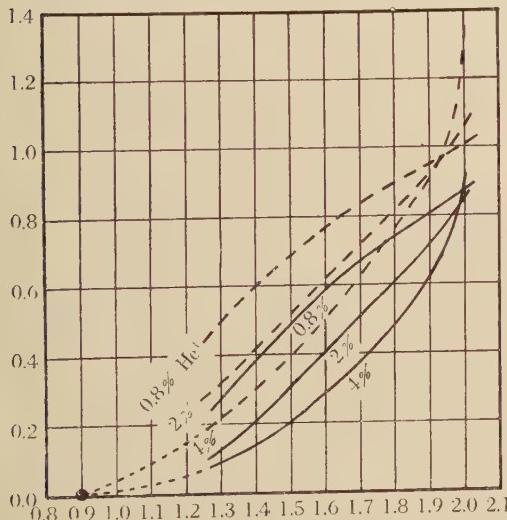

Fig. 1. Thermal Rayleigh disk torque τ in He^3-He^4 mixtures (relative to that for pure He^4) plotted as a function of temperature. Solid curves: the modified Bose-Einstein liquid theory. Dashed curves: Koide and Usui's theory. The circle at 0.9°K indicates the experimental value of Weinstock and Pellam for a 4% solution.

tion, except at temperatures close to the lambda-point of the mixture. Thus, as Weinstock and Pellam[3] suggested, the marked decrease in torque could be well explained without assuming that the He^3 does not participate in the normal fluid motion. The only published experimental work on this subject is that of Weinstock and Pellam[2]), which showed that the torque for a 4% solution was a few tenths percent of that for pure He^4 at 0.9°K. Both theories seem to be consistent with this result, though neither theory disregarding the effect of phonon excitation would be exact at this temperature. Experiments at higher temperatures are desirable in order to check the theory.

1) S. Koide and T. Usui, Prog. Theor. Phys. **6** (1951), 622.
2) Z. Mikura, Prog. Theor. Phys. **11** (1954), 25.
3) B. Weinstock and J. R. Pellam, Phys. Rev. **89** (1953), 521.

Composite States of Two Fermions through Fermi-type Interaction

Hiroshi Katsumori

Department of Physics,
Osaka Gakugei University

May 11, 1954

The suggestions have been made by several authors that bosons are composite, consisting of two fermions closely bound to each other. In this note it is attempted to consider the theory of Fermi and Yang[1] from the standpoint of a relativistic bound state problem.[2],[3]* However, the homogeneous Bethe-Salpeter equation for this problem has no solution as far as we assume that two fermions interact with each other by means of a contact interaction which leads to divergent integrals. Then a non-local form factor[5] with 4 end points is introduced as a resort to cut off the momentum relativistically.

Thus the Bethe-Salpeter equation in the ladder approximation becomes

---

* Lately Prof. S. Watanabe[4] has presented his fusion theory which deals with the similar subject from a different point of view.

$$\psi(1, 2) = ig \int K_+(1, 3') K_+^{\cdot}(2, 4') \varLambda F(3', 4', 1', 2')$$

$$\times \psi(1', 2') d(1', 2', 3', 4'), \tag{1}$$

where ψ stands for an amplitude with 16 components, K_+ a positron theoretical propagation function, g a coupling constant, and $\varLambda$ a certain set of Dirac matrices. The form factor with 4 end points is defined after the work of Møller and Kristensen, and its Fourier representation can be written as

$$F(x_1, x_2, x_3, x_4) = (2\pi)^{-16} \int F(f_1, f_2, f_3, f_4)$$

$$\cdot \exp\{-i[f_1 \cdot x_1 + f_2 \cdot x_2 + f_3 \cdot x_3 + f_4 \cdot x_4]\} df_1 df_2 df_3 df_4$$

$$= (2\pi)^{-12} \int G(P, f, f')$$

$$\cdot \exp\{-i[P \cdot (X - X') + f \cdot x + f' \cdot x']\} dP df df', \tag{2}$$

with

$$a \cdot b = a_\mu b_\mu = a_4 b_4 - \vec{a} \cdot \vec{b},$$

$$X = (x_1 + x_2)/2, \qquad X' = (x_3 + x_4)/2,$$

$$x = x_1 - x_2, \qquad x' = x_3 - x_4,$$

and $P = f_1 + f_2 = -(f_3 + f_4), \qquad f = (f_1 - f_2)/2,$

$$f' = (f_3 - f_4)/2.$$

Then we assume a G-factor with 4 end points in the following;

$$G(P, f, f') = G(\varPi^2, \varPi'^2) = 1 \quad \text{(both } \varPi^2 \text{ and } \varPi'^2 < K^2)$$

$$= 0 \quad \text{(otherwise)}, \tag{3}$$

where $\varPi^2 = -f_\mu^2 + (P_\mu f_\mu)^2/P_\mu^2$, which reduces to $\vec{f}^2$ in the rest system of the total system ($\vec{P} = 0$) and K means the cut-off momentum. If we set

$$G(\varPi^2, \varPi'^2) = G(\varPi^2) G(\varPi'^2) \tag{4}$$

using G-factors with 3 end points, the property assumed in (3) is satisfied. Inserting the relations (2), (4) and the Fourier representation of K_+ into (1) and putting $\psi(x_1, x_2) = \exp(-i P_\mu X_\mu) \varphi(x_\mu)$ where $(P_\mu P_\mu)^{1/2} = P$ is the rest mass of the total system, we get the following equation for $\varphi(x)$

$$\varphi(x) = -(2\pi)^{-8} ig$$

$$\times \int \frac{G(\varPi^2) e^{-iP \cdot x} df}{[\gamma_\mu^{(1)} (\tfrac{1}{2} P_\mu + f_\mu) - m][\gamma_\mu^{(2)}(\tfrac{1}{2} P_\mu - f_\mu) - m]} \varLambda$$

$$\times \int G(\varPi'^2) e^{-iP' \cdot x'} \varphi(x') df' dx'. \tag{5}$$

Since $\int G(\varPi^2) e^{-iP \cdot x} \varphi(x) df dx = \varphi$ is a constant with 16 spinor components, equation (5) gives the behavior of $\varphi(x)$. In order to obtain the relation between

the eigenvalue of equation (5) (proportional to the coupling constant g) and the rest mass of the bound state P, we multiply both sides of (5) by $G(\varPi^2) \cdot \exp(-ip \cdot x)$ and integrate over p and x. Thus we get

$$\varphi = -(2\pi)^{-5} ig$$

$$\times \int \frac{G(\varPi^2) G(\varPi'^2) e^{-i(p+p') \cdot x} df df' dx}{[\gamma_\mu^{(1)}(\tfrac{1}{2} P_\mu + f_\mu) - m][\gamma_\mu^{(2)}(\tfrac{1}{2} P_\mu - f_\mu) - m]}$$

$$\cdot \varLambda \varphi. \tag{6}$$

Now, rewriting this equation in the rest system, integrating over x and f', and using the relation $[G(\vec{f}^2)]^2 = G(\vec{f}^2)$, we have

$$\varphi = -(2\pi)^{-5} ig \int \frac{G(\vec{f}^2) e^{-i(p_4 + p_4') t}}{[\gamma_4^{(1)}(\tfrac{1}{2} P + f_4') + \vec{\gamma}^{(1)} \cdot \vec{f} - m]}$$

$$\frac{df_4 df_4' dt df}{[\gamma_4^{(2)}(\tfrac{1}{2} P - f_4') - \vec{\gamma}^{(2)} \cdot \vec{f} - m]} \varLambda \varphi. \tag{7}$$

Putting

$$\varphi = \begin{pmatrix} \varphi_{11} & \varphi_{12} & \varphi_{13} & \varphi_{14} \\ \varphi_{21} & \varphi_{22} & \varphi_{23} & \varphi_{24} \\ \varphi_{31} & \varphi_{32} & \varphi_{33} & \varphi_{34} \\ \varphi_{41} & \varphi_{42} & \varphi_{43} & \varphi_{44} \end{pmatrix} = \begin{pmatrix} \varphi_1 & \varphi_2 \\ \varphi_3 & \varphi_4 \end{pmatrix}$$

and carrying out the integration over f_4, f_4', t and $\vec{p}$, we obtain separately two sets of the simultaneous equatoins; one of which regards to φ_1 and φ_4, and another to φ_2 and φ_3. For example, in case of *scalar coupling* ($\varLambda = 1$) we have

$$\lambda \varphi_1 = \{[(P/2m) + 1]^2 A + B\} \varphi_1 + \tfrac{1}{3} \nu C \varphi_4 \bigg\}, \tag{8}$$

$$\lambda \varphi_4 = \{[(P/2m) - 1]^2 A + B\} \varphi_4 + \tfrac{1}{3} \nu C \varphi_1 \bigg\}$$

$$\lambda \varphi_2 = -C \varphi_2 - \tfrac{1}{3} \nu C \varphi_3 \bigg\}, \tag{8'}$$

$$\lambda \varphi_3 = -C \varphi_3 - \tfrac{1}{3} \nu C \varphi_2 \bigg\}$$

where

$$\lambda = \frac{2(2\pi)^2}{m^2} \frac{1}{g}$$

$$A = \log[(K + \sqrt{K^2 + m^2})/m] - (2W/P)$$

$$\cdot \tan^{-1}[(P/2W)(K/\sqrt{K^2 + m^2})],$$

$$B = \tfrac{1}{2}\{(K\sqrt{K^2 + m^2})/m^2 \tag{9}$$

$$- \log[(K + \sqrt{K^2 + m^2})/m]\},$$

$$C = B - (W/m)^2 A, \quad W = \sqrt{m^2 - (P/2)},$$

$$\nu = (\vec{\sigma}^{(1)}, \vec{\sigma}^{(2)}).$$
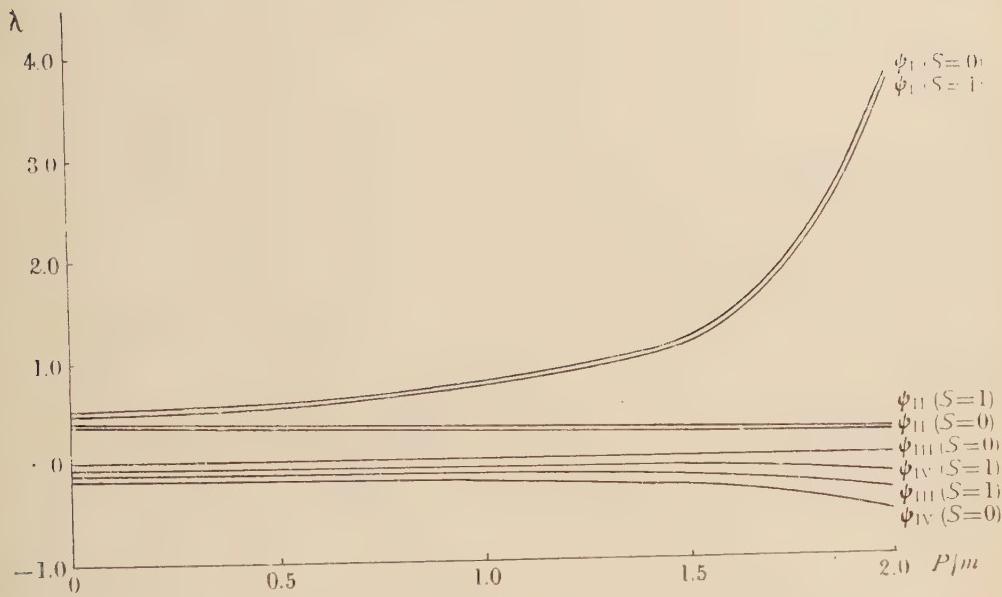
Fig. 1

Then, we can determine the eigenvalues and the eigenfunctions as follows. From (8) we have

$$\lambda = [(P/2m)^2+1]A+B\pm\sqrt{(P/m)^2A^2+(\tfrac{1}{3}\nu C)^2}$$

$$\equiv \lambda_I,\ \lambda_{II}, \qquad (10)$$

$$\begin{pmatrix} \varphi_1 & 0 \\ 0 & \varphi_4 \end{pmatrix} = f_1 \begin{pmatrix} I & 0 \\ 0 & -aI \end{pmatrix} \equiv \Psi_I,$$

$$f_1 \begin{pmatrix} I & 0 \\ 0 & a^{-1}I \end{pmatrix} \equiv \Psi_{II}, \qquad (11)$$

where $a = -\tfrac{1}{3}\nu C/[(P/m)A+\sqrt{(P/m)^2A^2+(\tfrac{1}{3}\nu C)^2}]$, which tends to unity in the limit of a local interaction ($K\to\infty$). And from (8′)

$$\lambda = -(1\pm\tfrac{1}{3}\nu)C \equiv \lambda_{III},\ \lambda_{IV}, \qquad (10')$$

$$\begin{pmatrix} 0 & \varphi_2 \\ \varphi_3 & 0 \end{pmatrix} = f_2 \begin{pmatrix} 0 & I \\ \mp I & 0 \end{pmatrix} \equiv \Psi_{III},\ \Psi_{IV}. \qquad (11')$$

In cases of other couplings we may solve equation (7) similarly;

pseudoscalar coupling ($A=\gamma_5^{(1)}\gamma_5^{(2)}$)

$$\lambda_I,\ \lambda_{II} = \tfrac{1}{3}\nu C$$

$$\pm \sqrt{(W/m)^4A^2+B^2+2[(P/2m)^2+1]A,B} \quad\Big\} \quad (12)$$

$$\lambda_{III},\ \lambda_{IV} = -C(\tfrac{1}{3}\nu\pm1)$$

$$\Psi_I,\ \Psi_{II} = g_1 \begin{pmatrix} I & 0 \\ 0 & \mp\beta I \end{pmatrix}, \quad \Bigg\}$$

$$\Psi_{III},\ \Psi_{IV} = g_2 \begin{pmatrix} 0 & I \\ \mp I & 0 \end{pmatrix}. \qquad (13)$$

Here $\beta = \dfrac{\{[(P/2m)-1]^2A+B\}}{\sqrt{(W/m)^4A^2+B^2+2[(P/2m)^2+1]AB}}$, which tends to unity in the limit of a local interaction.

vector coupling ($A=\gamma_\mu^{(1)}\gamma_\mu^{(2)}$)

$$\lambda_I,\ \lambda_{II} = [(P/2m)^2+1]A+B+(\nu/3)C\pm\sqrt{\Delta} \quad\Big\} \quad (14)$$

$$\lambda_{III},\ \lambda_{IV} = [(1+(\nu/3))\pm4\nu/3]C,$$

with

$$\Delta = (1-\nu^2)(P/m)^2A^2+(3/\nu)^2\{[(P/m)^2+2]A+4B\}^2,$$

$$\Psi_I = h_1 \begin{pmatrix} I & 0 \\ 0 & \varepsilon I \end{pmatrix}, \qquad \Psi_{II} = h_1 \begin{pmatrix} I & 0 \\ 0 & \varepsilon'I \end{pmatrix},$$

$$\Psi_{III},\ \Psi_{IV} = h_2 \begin{pmatrix} 0 & I \\ \mp I & 0 \end{pmatrix}. \qquad (15)$$

Here, $\varepsilon,\ \varepsilon' = \{-(P/m)A$

$$\pm\sqrt{(1-\nu^2)(P/m)^2A^2+(\nu/3)^2[((P/m)^2+2)A+4B]^2}$$

$\times 1/\nu\{(P/m)A+\tfrac{1}{9}[((P/m)^2+2)A+4B]\}\}$, which tend

to $\pm\dfrac{\nu}{|\nu|} = \begin{cases} \mp1 & (S=0) \\ \pm1 & (S=1) \end{cases}$ in the limit of a local

interaction.
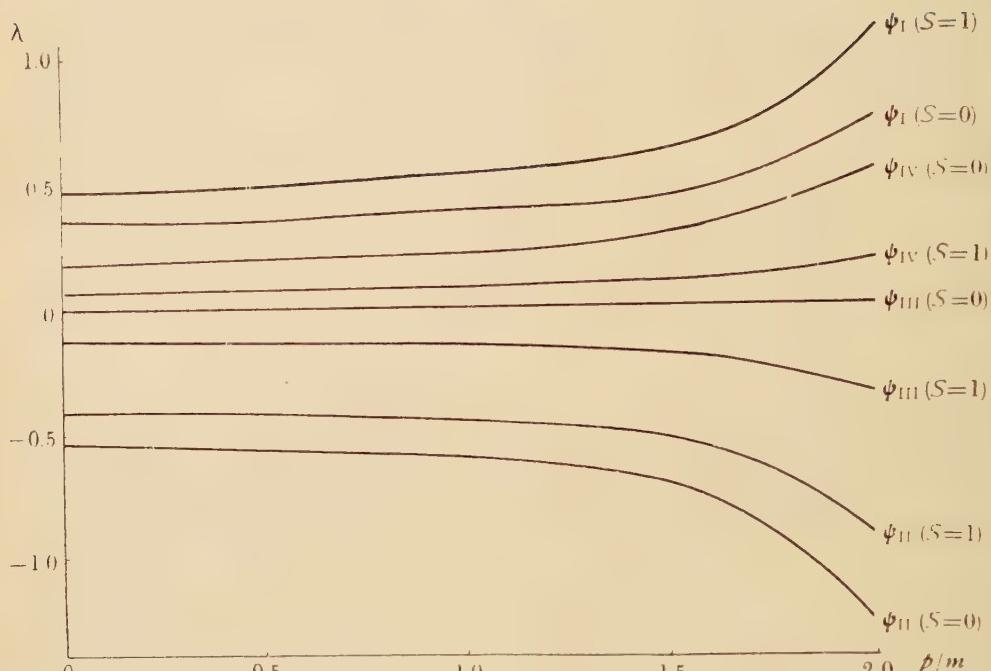
Fig. 2
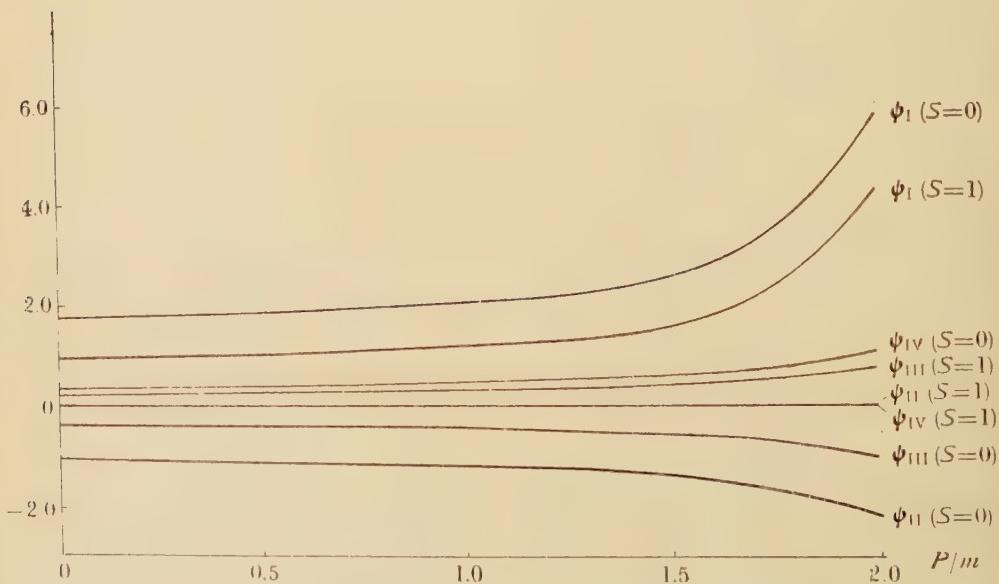
Fig. 3

According to the above results, the transformation properties of these eigenfunctions under space inversion and time reflection are tabulated in Tables 1, and 2. There e+ (0), for example, means that the corresponding eigenfunction consists of terms with even and odd character under the transformation concerned and the odd part vanished in a local limit.

The relations between $\lambda(\backslash 1/g)$ and P/m are shown in Figures 1, 2, and 3, which are drawn for $K=m$. From these figures we may observe the following remarks.

(i) We have the solutions of 4 singlet states and of 4 triplet states, and these solutions are distributed to both signs of the coupling constant irrespective of the coupling types and of the cut-off momentum. This fact seems to show that a vector coupling is not preferable as well as other types of coupling, against the theory of Fermi and Yang, since this coupling also can not always forbid a proton and a neutron to be directly combined.

(ii) A suitably chosen value of g corresponds to several (at most 4) different values of P/m. For example, in Fig. 1, assuming m as the mass of a nucleon and choosing $\lambda=0.531$, we have $P=276m_e$ (Ψ_I, $S=0$) and $735m_e$ (Ψ_I, $S=1$), or for the case $\lambda=0.352$, $P=276m_e$ (Ψ_{II}, $S=0$) and $1102m_e$ (Ψ_{II}, $S=1$) where m_e denotes the mass of a electron. Of course, it is meaningless to compare these mass spectra with the bosons in existence, because these numerical values are fairly sensitive to the magnitude of K.

We wish to express our most gratitude to Prof. C. Hayashi and Mr. S. Tani for their helpful guidance and valuable advice.

1) E. Fermi and C. N. Yang, Phys. Rev. **76** (1949), 1739.

2) C. Hayashi, and Y. Munakata, Prog. Theor. Phys. **7** (1952), 481.

3) C. Hayashi, Annual meeting of the Physical Society of Japan, Oct. 1952.

4) S. Watanabe, Phys. Rev. **91** (1953), 771.

5) P. Kristensen and C. Møller, Dan. Mat. Fys. Medd., **27** (1952), no. 7.

Table 1. Scalar coupling and pseudoscalar coupling

	Ψ_I	Ψ_{II}	Ψ_{III}	Ψ_{IV}
time inversion	o+(e)	e+(o)	e	o
space inversion	e+(o)	o+(e)	e	o
space and time inversion	o	o	e	e

Table 2. Vector coupling

		Ψ_I	Ψ_{II}	Ψ_{III}	Ψ_{IV}
time inversion	$S=0$	o+(e)	e+(o)	e	o
	$S=1$	e+(o)	o	e	o
space inversion	$S=0$	e+(o)	o+(e)	e	o
	$S=1$	o+(e)	e	e	o
space and time inversion	$S=0$	o	o	e	e
	$S=1$	o	o	e	e

Configuration Mixing and Magnetic Moments of Nuclei

Akito Arima and Hisashi Horie

Department of Physics, University of Tokyo

May 24, 1954

It is well-known that many properties of nuclei can be interpreted quite successfully by the shell model with strong spin-orbit interaction.[1],[2] However, some assumptions for this model have been derived as the consequences of short-range interaction between nucleons.[3] It is therefore expected that the different nucleon configurations should inevitably get mixed each other to a considerable extent in the stationary states of nuclei. The necessity of mixed configuration in nuclei was pointed out in the treatments of the first excited states in even-even nuclei and beta-transitions with anomalous ft values.[4],[5] Blin-Stoyl attempted the application of mixed configurations for the spin 1/2 nuclei under special assumptions.[6] Here, we also calculate the deviation of magnetic moments from Schmidt limits for more general case based on the configuration mixing.

For the sake of simplicity, we consider the nucleus with single proton or neutron in the outermost orbit j. There exist, in such a nucleus, also several numbers of orbits which are occupied by even numbers of nucleons coupled so as to make the total angular momentum zero, some of these orbits being not necessarily closed. We estimate the effects of excitation of these even numbers of nucleons caused by interaction with the single outermost orbit. The effects must approximately be additive for each orbit with even number of nucleons. Therefore, we represent

schematically the state of the nucleus, when the interaction between the single nucleon j and even number of nucleons in a certain orbit j_1 are ignored, as $|(j_1{}^n)_0 j j m\rangle$ (n: even), where the last j and m denote the total angular momentum of this system and its z-component. We shall next find the configurations which can mix with the above configuration and contribute to the magnetic moment linearly in the coefficient of mixing. In view of the property of magnetic moment operator, responsible configurations are evidently $|(j_1{}^{n-1}) j_1 (j_2 j)_J j m\rangle$, where the orbit j_2 must have the same orbital angular momentum as the orbit j_1, say l_1, and J is restricted by the relation $|j_1-j| \leq J \leq j_1+j$. Hence, the mixed configuration which we are interested in is expressed as

$$|(j_1{}^n)_0 j j m\rangle + \sum_J a_J (j_1{}^{n-1}) j_1 (j_2 j)_J j m\rangle. \quad (1)$$

We neglect the reduction of the amplitude of the first term in (1), assuming that all the coefficients of mixing of the configurations including those not contributing to the magnetic moment, are small. a_J's are the coefficients of mixing to be determined in the following and their magnitudes are found to be of the order 0.1 or less.

The correction to the magnetic moment of the usual single particle model can easily be obtained in a straightforward way while the term linear in a_J is given by

$$\delta\mu = \sqrt{n} \sum_J 2 a_J (g_s - g_l) [j / \sqrt{3(j+1)(2l+1)}]^{1/2}$$
$$\cdot \langle j_1, j_2 j(J) j | j_1 j_2 (1) j j \rangle, \quad (2)$$

where $g_s - g_l = 4.585$ and -3.826 n.m. for proton and neutron, respectively, and the last factor in (2) denotes the transformation coefficient between two coupling schemes of three angular momenta j_1, j_2 and j coupled to total angular momentum j. On the other hand, a_J can be evaluated by the simple perturbation theory. If we put zeroth-order energy differences between each of the second and the first configuration in (1) ΔE_J, it is well-known that

$$a_J = -\langle (j_1{}^n)_0 j j m | V | (j_1{}^{n-1}) j_1 (j_2 j)_J j m \rangle / \Delta E_J, \quad (3)$$

where V is the internucleonic interaction. Inserting the values of a_J into (2), we get the deviation of magnetic moment from the single particle value (Schmidt limit) due to the mixing of the configurations described by the second term in (1). If we adopt for the interaction V the δ-function type as the short-range limit and substitute the average value of the energy differences $\langle \Delta E \rangle$ so as to carry out

the summation with respect to J, the result becomes much simpler. The formulae for $\delta\mu$ is given in Table I. V_s and V_t are the interaction strengths in singlet and triplet states of two nucleon, respectively, and will be negative (*i.e.* attractive) and $|V_t| > |V_s|$ (*i.e.*, the attractive force in the triplet state is stronger than in the singlet state). And I is an integral to which generalized Slater integrals are reduced on account of δ-function type interaction:

$I = \frac{1}{2} \int_0^\infty [R_l(r) R_{l_1}(r)]^2 r^2 dr.$ Thus, we see from Table I that the theoretical deviation of the magnetic moments from the Schmidt limits have proper signs for each case in accord with the empirical data.[7] It is noticed that there are no corrections to $p_{1/2}$ states in our approximation while experimentally the deviations of magnetic moments of 1/2 (odd) nuclei from the Schmidt values are very small. Although we have assumed thus far that the orbit j_2 is empty in the non-interaction configuration, it is easy to modify our results to include the case in which even number of nucleons in j_2-orbit are present, coupled to zero angular momentum. The modification when the orbit j_2 coincides with orbit j is also at hand. Under such modifications, the signs do not alter so that above conclusion remains valid.

In order to compare the calculated deviation of magnetic moments from the Schmidt limits with the experimental values, we take the values of interaction constants, the integral I and the energy differences as

$$V_t = 1.5 V_s, \quad V_s I = -25/A \text{ Mev.}, \quad \text{and}$$
$$\langle \Delta E \rangle \propto (2l_1 + 1) A^{-2/3}. \quad (4)$$

Here, the values of $V_s I$ was determined in accordance with that deduced in the argument of pairing energy[3] and we assume that it depends upon mass number A only in order to avoid complexity. On the other hand, ΔE is the sum of the energy difference between single particle levels j_1 and j_2, and the average value of the differences between the pairing energy $(j_1{}^2)_0$ and the energy of $(j_2 j)_J$ states. Since the estimate of the latter is rather complicated and its value is expected to be positive, we adopted as $\langle \Delta E \rangle$ somewhat enlarged value of the former, *i.e.*, doublet splitting of the levels with orbital angular momentum l_1. Thus, the proportional constant of the expression of $\langle \Delta E \rangle$ is chosen to fit $\langle \Delta E \rangle = 5$ Mev for O^{17}. The calculated deviation of magnetic moments from the Schmidt limits are shown in Table II for some nuclei. The calculated values show good agreement with experimental values and it is remarkable that the deviation of magnetic moments from the schmidt limits

can be accounted for in such a simple way based upon mixed configuration in nuclei. Detailed account will appear soon in this journal.

1) M. G. Mayer, Phys. Rev. **78** (1950), 16.
2) O. Haxel, J. Jensen and H. Suess, ZS. f. Phys. **128** (1950), 295.
3) M. G. Mayer, Phys. Rev. **78** (1950), 22.
4) P. Stählin and P. Preiswerk, Nuov. Cim. **10** (1953), 1219.
5) A. de-Shalit and M. Goldhaber, Phys. Rev. **92** (1953), 1211.
6) R. J. Blin-Stoyle, Proc. Phys. Soc. **66A** (1953), 1158.
7) P.A.F. Klimkenberg, Rev. Mod. Phys. **24** (1952), 63.

Table I. Deviation of magnetic moments from the Schmidt limits in $n.\ m.$

		δ_μ	Contribution from Even Number of
Odd Proton Nuclei	$j=l+\frac{1}{2}$	$-\dfrac{n(l+2)l_1}{(2l+3)(2l_1+1)}\times\begin{cases}(-V_sl)\times 4.585/\langle \Delta E\rangle\\ \frac{1}{2}(V_s-V_t)l\times 3.826/\langle \Delta E\rangle\end{cases}$	Protons Neutrons
	$j=l-\frac{1}{2}$	$\dfrac{n(l-1)l_1}{(2l+1)(2l_1+1)}\times\begin{cases}(-V_sl)\times 4.585/\langle \Delta E\rangle\\ \frac{1}{2}(V_s-V_t)l\times 3.826/\langle \Delta E\rangle\end{cases}$	Protons Neutrons
Odd Neutron Nuclei	$j=l+\frac{1}{2}$	$\dfrac{n(l+2)l_1}{(2l+3)(2l_1+1)}\times\begin{cases}\frac{1}{2}(V_s-V_t)l\times 4.585/\langle \Delta E\rangle\\ (-V_sl)\times 3.826/\langle \Delta E\rangle\end{cases}$	Protons Neutrons
	$j=l-\frac{1}{2}$	$-\dfrac{n(l-1)l_1}{(2l+1)(2l_1+1)}\times\begin{cases}\frac{1}{2}(V_s-V_t)l\times 4.585/\langle \Delta E\rangle\\ (-V_sl)\times 3.826/\langle \Delta E\rangle\end{cases}$	Protons Neutrons

Table II. Comparison of the calculated and experimental values of deviation of magnetic moments from the Schmidt limits in $n.\ m.$

Nucleus	Spin	P-conf.*	N-conf.*	$\delta\mu_{cal.}$	$\delta\mu_{exp.}$
$_9F_{10}^{19}$	$\frac{1}{2}$	$s\frac{1}{2}$	$(d\frac{5}{2})^2$	-0.13	-0.16
$_{15}P_{16}^{31}$	$\frac{1}{2}$	$(d\frac{5}{2})^6s\frac{1}{2}$	$(d\frac{5}{2})^6$	-1.43	-1.66
$_{29}Cu_{34}^{63}$	$\frac{3}{2}$	$(f\frac{7}{2})^8p\frac{3}{2}$	$(f\frac{7}{2})^8(f\frac{5}{2})^2$	-1.42	-1.56
$_{29}Cu_{36}^{65}$	$\frac{3}{2}$	,,	$(f\frac{7}{2})^8(f\frac{5}{2})^4$	-1.29	-1.41
$_{17}Cl_{18}^{35}$	$\frac{3}{2}$	$(d\frac{5}{2})^6d\frac{3}{2}$	$(d\frac{5}{2})^6(d\frac{3}{2})^2$	0.35	0.70
$_{17}Cl_{20}^{37}$	$\frac{3}{2}$	,,	—	0.30	0.56
$_{51}Sb_{70}^{121}$	$\frac{5}{2}$	$(g\frac{9}{2})^{10}d\frac{5}{2}$	$(d\frac{5}{2})^6(h\frac{11}{2})^6$	-1.30	-1.43
$_{23}Bi_{136}^{209}$	$\frac{9}{2}$	$(h\frac{11}{2})^{12}h\frac{9}{2}$	$(i\frac{13}{2})^{14}$	0.81	1.46
$_{14}Si_{19}^{29}$	$\frac{1}{2}$	$(d\frac{5}{2})^6$	$(d\frac{5}{2})^6s\frac{1}{2}$	1.37	1.35
$_{50}Sn_{65}^{115}$	$\frac{1}{2}$	$(g\frac{7}{2})^{10}$	$(d\frac{5}{2})^6s\frac{1}{2}$	1.18	1.00
$_{24}Cr_{29}^{53}$	$\frac{3}{2}$	$(f\frac{7}{2})^4$	$(f\frac{7}{2})^8f\frac{5}{2}$	1.42	1.44
$_{16}S_{17}^{33}$	$\frac{3}{2}$	,,	$(d\frac{5}{2})^6d\frac{3}{2}$	-0.36	-0.51
$_8C_9^{17}$	$\frac{5}{2}$	——	$d\frac{5}{2}$	0	0.02
$_{40}Zr_{51}^{91}$	$\frac{5}{2}$	——	$(g\frac{9}{2})^{10}d\frac{5}{2}$	1.07	0.9
$_{60}Nd_{83}^{143}$	$\frac{7}{2}$	$(d\frac{5}{2})^2$	$(h\frac{11}{2})^{12}(h\frac{9}{2})^2f\frac{7}{2}$	1.07	0.91

The P- and N-conf. denote only the proton and neutron orbits which affect to the magnetic moments, respectively.

Interpretation of the Excitation of Gamma-rays by 10 Mev Neutrons

Satio HAYAKAWA

Research Institute for Fundamental Physics, Kyoto University

and

Ken KIKUCHI

Department of Physics, Osaka University

(Received March 25, 1954)

Gamma-rays from medium weight and heavy nuclei excited by 10 Mev neutrons are calculated on the basis of the compound nucleus formalism. Calculated gamma-ray yield shows a fair agreement with the experimental data by Wakatsuki and Sugimoto, if the empirical temperature-mass number relation is adopted. The competition of various emission processes is found to be important to account for the experiment.

§ 1. Introduction

The purpose of this paper is to investigate the A-dependence of the excitation of gamma-rays by fast neutrons. Wakatsuki and Sugimoto[1] observed fluorescent gamma-rays from 32 elements excited by Li-D neutrons. Observed gamma-ray intensities increase almost linearly with mass number A up to $A \cong 120$ and then become nearly constant or slightly decrease at heavier nuclei. They concluded that this trend of their experimental results was not in agreement with the gamma-ray intensity expected from Weisskopf's statistical theory.[*] Miyatake[2] analysed these results on the basis of the core structure for heavy nuclei.

We do not think, however, that the statistical description of Weisskopf fails to account for this experiment. But the gamma-ray yield may be explained by a temperature of an appropriate A-dependence. We are able to calculate the gamma-ray yield without assuming any specific nuclear model, since the A-dependence of the temperature is known from other data, though the A-dependence should be based upon a specific nuclear structure.

In this paper the gamma-ray intensity excited by 10 Mev neutrons is calculated along the compound nucleus formalism. In order to carry out our calculations we have to consider succesively : (i) the determination of the separation energies, (ii) the determination of level densities, (iii) the determination of the cross section for the formation of the compound nucleus.

---

* Yamabe and Sanada (Prog. Theor. Phys. **5** (1950), 1049) argued that the irregularity in the gamma ray yields from various nuclei was due to the shell structure. In our phenomenological treatment, however, this effect by the shell structure enters as the fluctuation of binding energies.

§ 2.　Determination of necessary quantities

(1)　*Separation energies.*　Our attention is given to the overall A-dependence of gamma-ray intensity, that means, the small fluctuations of magnitudes of separation energies are ignored. In this case we may assume that the binding fraction f is a function of A only and that it is smooth enough so that its derivative exists. Then the separation energy S_n for neutron is given by[3]

$$S_n \cong f + (A-1) df/dA . \tag{1}$$

The values of S_n used in our calculation are shown in Fig. 1.

(ii)　*Level densities.*　We use the semi-empirical formula for the level density of a nucleus as a function of excitation energy E:

$$w(E) = c \exp(2\sqrt{aE}) \tag{2}$$

with

$$c_{\text{even } A} = \frac{1}{2} c_{\text{ odd } A}, \tag{3}$$

Fig. 1.　Neutron separation energy S_n and binding fraction f vs. mass number A. Solid line is S_n and broken-line is f (Mev).

where c and a are two parameters which have to be determined from the scant experimental data. Blatt and Weisskopf have given estimates for c and a for mass numbers around 27, 63, 115, 181, and 231.[4] We shall obtain c and a by interpolation of Blatt and Weisskopf's values. These values for odd A nuclei are shown in Fig. 2 and Fig. 3.
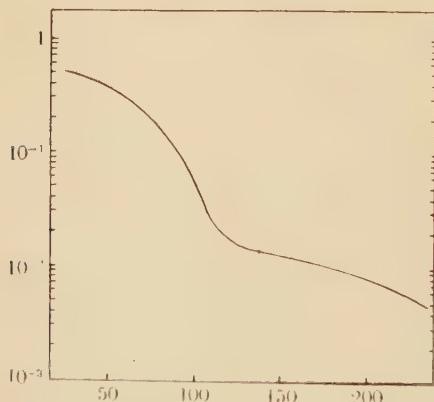

Fig. 2.　Values of c vs. mass number for odd A nuclei.
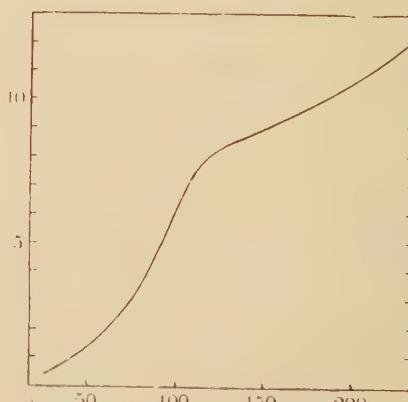
Fig. 3.　Values of a in Mev^{-1} vs. mass number.

(iii)　*Cross sections for the formation of a compound nucleus.*　It is necessary to know the cross section σ_c for neutrons and protons. For neutrons the formula of σ_c has been given by Feshbach and Weisskopf.[5] For protons the calculated values of σ_c are

listed by Blatt and Weisskopf[6] for Z larger than 10. We simply use them interpolated from the above table.

It must be noted, however, that σ_c for incident neutrons is slightly larger than those obtained from experimental data[7] for heavy nuclei as compared in Fig. 4. Hence we employ the experimental σ_c for the compound nucleus formation by incident neutrons. But for σ_c appearing in the particle width we are obliged to use the theoretical values. The difference due to this may not be so large as the approximations made in other parts.
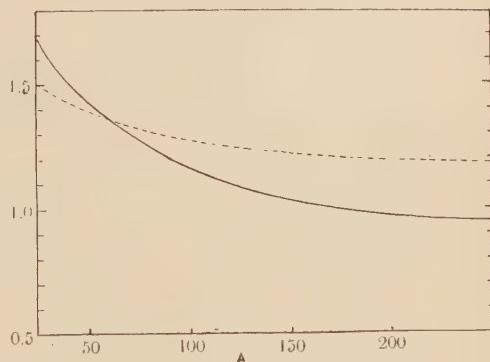
Fig. 4. Cross section for the formation of a compound nucleus by 14-Mev neutrons *vs.* mass number. Solid line is experimental value and broken-line is that calculated by Feshbach-Weisskopf's theory.

§ 3. Calculation of emission widths

According to the compound nucleus formalism, the probability of evaporating a particle i is given by

$$F_i/\sum_j F_j ,\qquad\qquad (4)$$

where F_i is a quantity proportional to the partial width for disintegration with emission of i and the sum is extended over all kinds of partial widths. It is shown from detailed balancing that, if i is a particle[8],

$$F_i(W_i) = \frac{2M_i}{\hbar^2} \int_0^{W_i} \epsilon\, \sigma_c^{(i)}(\epsilon)\, w_R(W_i-\epsilon)\, d\epsilon . \qquad\qquad (5)$$

In this formula W_i is the available energy, that is equal to the excitation energy E_{ex} minus the separation energy S_i of particle i; and $\sigma_c^{(i)}(\epsilon)$ is the σ_c for particle i with the kinetic energy ϵ. w_R is the level density of the residual nucleus and M_i is the mass of an emitted particle i. When i is a gamma-ray, the quantity F_i is given by[9]

$$F_\gamma(E_{ex}) \simeq \frac{3}{4}\frac{e^2}{\hbar c}\left(\frac{R}{\hbar c}\right)^2 \frac{1}{D_0}\int_0^{E_{ex}} \epsilon^3 w_R(E_{ex}-\epsilon)\, d\epsilon , \qquad\qquad (6)$$

where D_0 is the level distance of low-lying levels and it is approximately 0.5 Mev for medium weight and heavy nuclei, and R is the nuclear radius.

Formulas (4), (5) and (6) allow us to calculate the cross section of each type of reactions initiated by fast neutrons. The cross section of a $(n ; i, j)$ reaction can then be expressed *approximately* by

$$\sigma(n ; i, j) = \sigma_c^{(n)}(10\text{ Mev})\, (F_i/\sum_k F_k)\, (F_j/\sum_l F_l)' , \qquad\qquad (7)$$

where the prime means that this quantity is taken at the intermediate residual nucleus.

The emission of deuterons, tritons or α-particles may be negligibly small.   The emission of protons is also negligible at heavy nuclei.   The average excitation energy E_{ec}' of the intermediate residual nucleus is given by

$$E_{ec}' = E_{ec} - S_i - 2\theta', \tag{8}$$

where θ' is the temperature of the intermediate residual nucleus and is equal to $(E_{ec}'/a)^{\frac{1}{2}}$. When i is any particle, θ' is lower than when i is a gamma-ray.

§ 4 Results and discussions

We can now evaluate the cross sections.   Calculated results are shown in Fig. 5 and 6.

Once the compound nucleus is formed, competition sets in between different modes of decays.   Because of the existence of the Coulomb barrier, the emission of neutrons is predominant over that of charged particles. In heavy nuclei it becomes energetically possible that the compound nucleus can emit two neutrons.   This kind of reaction was not taken into account in the previous analysis[1)2], but its importance was recognized by Heidmann and Bethe[10] in their ananlysis of photonuclear reactions.   The emission of two neutrons leaves the residual nucleus in the ground state or low-lying levels, so that the gamma-ray energy from such a residual nucleus is very small.   Thus the occurrence of $(n ; 2n)$ reaction for nuclei with $A \gtrsim 170$ decreases the gamma-ray yield, which is proportional to the gamma-ray energy in the experiment referred.[1)]   Both, the $(n ; \gamma)$ and the $(n ; \gamma, n)$ cross sections increase with A, reach a maximum
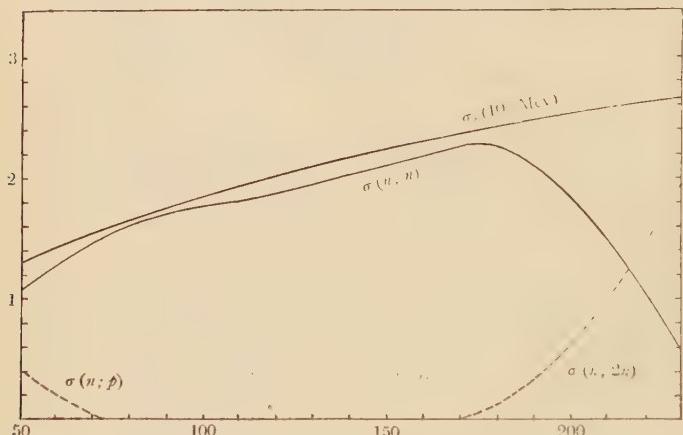
Fig. 5.   Cross sections of $(n ; n)$, $(n ; 2n)$ $(n ; p)$ reactions in barn $vs.$ mass number.
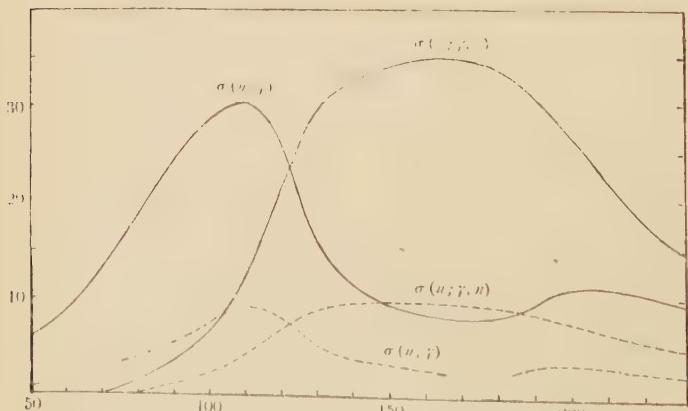

Fig 6.   Cross sections of $(n ; \gamma)$ and $(n ; \gamma, n)$ reactions in milibarn $vs.$ mass number.   Solid line refers to even A target nuclei and broken-line refers to odd A.

and then decrease for heavy nuclei.

A comparison can be made now between experimental results and our theoretical results. For the comparison it is necessary to obtain the relative gamma-ray yield which is given by the product of the cross section and the corresponding average gamma-energy. These are shown in Fig. 7.

Although we have ignored the detailed structure of the individual nucleus in the above treatment, a special attention is paid to ^{103}Ag, ^{121}Sb and ^{113}Cd for checking the influence by the fluctuation of separation energies. The gamma-ray yields are calculated using observed values of S_n for these nuclei, and the results are in good agreement with the experimental data as shown by double circles in Fig. 7.

The determination of level density parameters is not quite correct, because our information about level density is very scant and Blatt-Weisskopf's values are fitted to experiments around 1 Mev excitation, whereas the excitation energy is very high in our case.

Fig. 7.   Relative gamma-ray yield *vs.* mass number. Ordinate is arbitrary scale. Solid line and broken line indicate even *A* and odd *A* target nuclei respectively. Double circles are those calculated for ^{109}Ag, ^{114}Cd, and ^{121}Sb. Plotted experimental values are :
- •   Wakatsuki and Sugimoto : Reference 1)
- ○   H. Aoki : Proc. Phys. Math. Soc. Japan **18** (1937), 115.
- ×   D. E. Lea : Proc. Roy. Soc. A159 (1935), 637.

Aoki's data are those by 2.5 Mev *d-d* neutrons and Lea's data are by Po(α)-Be neutrons (of mean energy about 5 Mev).

However the general trend of *A*-dependence of the gamma-ray yield will be conserved even if other values[11] are employed.

Thus we may summarize that the following facts make the gamma-ray yield nearly constant for heavy nuclei ; (1) the occurrence of (n ; $2n$) reaction, (2) the decrease of (n ; γ) and (n ; γ, n) cross sections, and (3) the smaller magnitude of the actual inelastic cross section for incident neutrons than the theoretical values.

Thanks are due to Professors Wakatsuki and Miyatake for a number of discussions.

References

1) T. Wakatsuki and K. Sugimoto, Jour. Phys. Soc. Japan **7** (1952), 249.
2) O. Miyatake, Prog. Theor. Phys. **7** (1952), 285.
3) J. M. Blatt and V. F. Weisskopf, " *Theoretical Nuclear Physics* " John Wiley & Sons, New York, (1952), p. 6.
4) Reference 3) p. 372.
5) H. Feshbach and V. Weisskopf, Phys. Rev. **76** (1949), 1550.
6) Reference 3) p. 352.
7) D. Phillips, R. Davis and E. Graves, Phys. Rev. **88** (1952), 600.
8) Reference 3) p. 369.
9) Reference 3) p. 649.
10) J. Heidmann and H. A. Bethe, Phys. Rev. **84** (1950), 274.
11) P. C. Gugelot, Phys. Rev. **81** (1951), 51.
      E. Graves, and L. Rosen Phys. Rev. **89** (1953), 343.

A Method in Quantum Statistical Mechanics, III

—— *Virial Expansion for the Quantum Gas* ——

Hiroshi ICHIMURA

Physics Department, Tokyo Institute of Technology

Oh-Okayama, Meguro-ku, Tokyo

As an application of the general expansion formula of the grand partition function for the system of interacting particles, we derive the virial expansion formula for quantum gases under the assumption that the degeneracy parameter λ is very small. We get a formula for the second virial coefficient which is effective at extremely low temperatures.

§ 1. Virial expansion formula

As the second application of the formula $(1 \cdot 13)$ in Part I,[1] we now treat the case in which the assumption

$$\lambda \ll 1$$

is allowed. This corresponds usually to the gaseous phase, and in this case the equations

$$pV = \kappa T \sum_r \log(1 \pm \lambda e^{-\beta \varepsilon_r})^{\pm'} + \kappa T \sum_i J^i C_i,$$

$$N = \sum_r \frac{\lambda e^{-\beta \varepsilon_r}}{1 \pm \lambda e^{-\beta \varepsilon_r}} + \kappa T \sum J^i \lambda \frac{\partial}{\partial \lambda} C_i \qquad (1 \cdot 1)^*$$

may be used in the expanded form in powers of λ. If we stop with the term λ^i, we obtain the expansion up to the i-th virial coefficients, as may be seen easily.

Now we can assume the form

$$\sum J^i C_i = \lambda^2 V \tilde{C}_2 + \lambda^3 V \tilde{C}_3 + \cdots \qquad (1 \cdot 2)$$

from its definition and obtain

$$pV = \kappa T V \left(\sum_{i=1}^{\infty} (\lambda/i) F_i + \sum_{l=2}^{\infty} \lambda^l \tilde{C}_l \right) \qquad (1 \cdot 3)$$

and

$$N = V \left(\sum_{i=1}^{\infty} \lambda^i F_i + \sum_{l=2}^{\infty} l \lambda^l \tilde{C}_l \right), \qquad (1 \cdot 4)$$

where

---

* N is the average number of particles in volume V. Other notations are the same as in Part I.

$$VF_i = \sum_r e^{-i\beta\epsilon_r} 2\pi g V \left(\frac{2m}{h^2}\right)^{3/2} \int_0^\infty \sqrt{\mathcal{E}}\, e^{-\beta\epsilon}d\mathcal{E} \tag{1.5}$$

$$= Vg\Lambda^{3/2}(\pm1)^{i-1}/i^{3/2}, \tag{1.6}$$

$$\Lambda = 2\pi m\kappa T/h^2 \tag{1.7}$$

$$(+\text{Bose}, \ -\text{Fermi}).$$

Then we assume the expanded form of λ in powers of l^{-1} as in Mayer's theory of the imperfect gas and determine the coefficients successively.

$$\lambda = a^1/V + a^2/V^2 + a^3/V^3 + \cdots \tag{1.8}$$

Inserting this in (1.3) and (1.4), we get

$$pV = \kappa T\Big\{a_1F_1 + \frac{1}{V}\Big(a_2F_1 + \frac{a_1^2}{2}F_2 + a_1^2\tilde{C}_2\Big) +$$
$$+ (1/V^2)\,(a_3F_1 + a_1a_3F_2 + a_1^3F_3/3 + 2a_1a_2\tilde{C}_2 + a_1^3\tilde{C}_3)\Big\}$$
$$+ O(1/l^{3}) \tag{1.9}$$

and

$$N = a_1F_1 + (1/V)\,(a_2F_1 + a_1^2F_2 + 2a_1^2\tilde{C}_2) +$$
$$+ (1/V^2)\,(a_3F_1 + 2a_1a_2F_2 + a_1^3F_3 + 4a_1a_2\tilde{C}_2 + 3a_1^3\tilde{C}_3) +$$
$$+ O(1/l^{-3}). \tag{1.10}$$

From (1.10) we obtain

$$a_1 = N/F_1,$$
$$a_2 = (N^2/F_1^2)(-F_2\tilde{C}_2),$$
$$a_3 = (N^3/F_1^3)\Big(\frac{8}{F_1^2}\Big(\frac{F_2}{2}+\tilde{C}_2\Big)^2 - \frac{F_3}{F_1} - 3\frac{\tilde{C}_2}{F_1}\Big), \tag{1.11}$$

. . . .

Hence we get the virial expansion formula from (1.9)

$$pV/N\kappa T = 1 + B(T)/V + B(V)/V^2 + \cdots,$$

where

$$B(T) = B_0(T) + B_1(T), \tag{1.12}$$

$$B_0(T) = \frac{-NF_2}{2F_1^2} = \mp \frac{N}{4\sqrt{2}\,g}\Big(\frac{h^2}{2\pi m\kappa T}\Big)^{3/2}, \tag{1.13}$$

$$(-\text{Bose}, +\text{Fermi})$$

$$B_1(T) = -N\,\tilde{C}_2/F_1^2 = -N\tilde{C}_2/g^2\Lambda^3 \tag{1.14}$$

and

$$C(T) = \frac{N^2}{F_1^2}\left(\frac{4}{F_1}\left(\frac{1}{2}F_2 + \tilde{C}_2\right)^2 - \frac{2}{3}F_3 - 2\tilde{C}_3\right) \tag{1·15}$$

$$= C_0(T) + C_1(T),$$

$$C_0(T) = \left(\frac{1}{8} - \frac{2}{3^{5/2}}\right)\frac{N^2}{g^2}\left(\frac{h^2}{2\pi m\kappa T}\right)^3, \tag{1·16}$$

$$C_1(T) = -\frac{N^2}{g^2}\left(\frac{h^2}{2\pi m\kappa T}\right)^2\left(\pm\frac{\tilde{C}_2}{\sqrt{2}\,g}A^{-3/2} - 4\frac{(\tilde{C}_2)^2}{g^2}A^{-3} + 2\frac{\tilde{C}_3}{g}A^{-3/2}\right). \tag{1·17}$$

The quantity B_0 and C_0 are the virial coefficients for the ideal quantum gas. In similar fashions, we can go further to the higher terms, but the calculation of $\tilde{C}_i$ becomes increasingly very troublesome. So we calculate only the second virial coefficients in the next section.

§ 2. The second virial coefficient

The quantity $B_1(T)$ in $(1·14)$ shows the effect of the interparticle interaction on the second virial coefficient. To evaluate this term we must know the explicit form of $\tilde{C}_2$. We expand all terms in the series $\sum J^i C_i$ in powers of λ and get the following results.
From I–$(2·2)$ we get

$$JC_1 = -\beta\lambda^2\sum_{j\leq k}\sum e^{-\beta(\varepsilon_j+\varepsilon_k)}G_{jk} \pm \beta\lambda^3\sum_{j\leq k}\sum\{e^{-\beta(\varepsilon_j+2\varepsilon_k)} + e^{-\beta(2\varepsilon_j+\varepsilon_k)}\}G_{jk} + \cdots. \tag{2·1}$$

From I–$(2·5)$

$$J^2 D_2 = \frac{\lambda^2\beta^2}{2}\sum_{j\neq k\neq l}\sum\sum \exp[-\beta(\varepsilon_j + \varepsilon_k + \varepsilon_l)]G_{jk}G_{jl} + O(\lambda^3). \tag{2·2}$$

From I–$(2·7)$

$$J^3 D_3 = O(\lambda^3). \tag{2·3}$$

From I–$(2·15)$ and I–$(2·15)$

$$J\langle N_2\rangle = \lambda^2\beta^2 J^2 S_{pq,rs}\frac{(G_{pq,rs})^2 e^{-\beta(\varepsilon_r+\varepsilon_s)}}{\beta(\varepsilon_p + \varepsilon_q - \varepsilon_r - \varepsilon_s)} + O(\lambda^3). \tag{2·4}$$

And from the definition we know that we can set

$$J^3 \ll \langle F_2\rangle - \langle D\rangle\langle N_2\rangle \gg = \lambda^2 J^3\beta^3\bar{G}_3 + O(\lambda^3), \tag{2·5}$$

$$J^4 \ll \langle F_4\rangle (\cdots)\gg = \lambda^2 J^4\beta^4\bar{G}_4 + O(\lambda^3), \tag{2·6}$$

etc.
So we obtain

$$\sum J^i C_i = \lambda^2(-\beta\bar{G}_1 + \beta^2 J^2\bar{G}_2 - \beta^3 J^3\bar{G}_3 + \cdots) \tag{2·7}$$

and then

$$\tilde{C}_2 = -\beta J \frac{\overline{G}_1}{V} + \beta^2 J^2 \frac{\overline{G}_2}{V} - \beta^3 J^3 \frac{\overline{G}_3}{V} + \cdots , \tag{2·8}$$

where

$$\beta J \overline{G}_1 = \beta J \sum_{j \leq k} \sum e^{-\beta(\varepsilon_j + \varepsilon_k)} G_{jk} , \tag{2·9}$$

$$\beta^2 J^2 \overline{G}_2 = \beta^2 J^2 S_{pq,rs} \frac{(G_{pq,rs})^2}{\beta(\varepsilon_p + \varepsilon_q - \varepsilon_r - \varepsilon_s)} e^{-\beta(\varepsilon_r + \varepsilon_s)} . \tag{2·10}$$

We stop with $\overline{G}_2$ for the moment.

(1) *The evaluation of* $\overline{G}_1$.

Now we must give the actual form of the potential function $G(r)$. We cannot use the usual Lennard-Jones type functions on account of its singular nature at the origin. We adopt a function which is the sum of the terms whose form is Ae^{-ar}/r, that is

$$JG(r) = J \sum_a \frac{A}{r} e^{-ar}, \tag{2·11}$$

here we restrict in two terms

$$JG(r) = (J/r) (Ae^{-ar} - Be^{-br}). \tag{2·11'}$$

This is a modification of the Morse potential and we determine the constants JA, a etc. in such a way that $(2·11)$ may reproduce the actual potential function as precisely as possible.

With the above form of $G(r)$, we obtain

$$J\overline{G}_1 = \frac{1}{2} \sum_a \left(\frac{4\pi JA}{Va^2} \right) \left(\frac{gV}{(2\pi)^3} \right)^2 \iint \left(1 \pm \frac{a^2}{g(a^2 + k_{jk}^2)} \right) e^{-\beta(\varepsilon_j + \varepsilon_k)} d\mathbf{k}_j d\mathbf{k}_k , \tag{2·12}$$

where

$$JG_{jk} = \frac{1}{V} \sum_a \left(\frac{4\pi JA}{a^2} \right) \left(1 \pm \left(\frac{a^2}{g(a^2 + k_{jk}^2)} \right) \right) \tag{2·13}$$

$$(+\text{Bose}, \ -\text{Fermi})$$

has been used (Spin summation has been performed). This can be transformed as

$$\frac{J\overline{G}_1}{V} = \frac{1}{2} \frac{g^2}{(2\pi)^6} \sum_a \left(\frac{4\pi JA}{a^2} \right) \left\{ 1 \pm \frac{1}{g} K - (\pm) \frac{1}{g} K_a \right\} \tag{2·14}$$

$$(+\text{Bose}, \ -\text{Fermi})$$

where

$$K = \iint e^{-\beta(\varepsilon_j + \varepsilon_k)} d\mathbf{k}_j d\mathbf{k}_k = (2\pi)^6 / \Lambda^3 , \tag{2·15}$$

$$K_a = \iint e^{-\beta(\varepsilon_j + \varepsilon_k)} \frac{k_{jk}^2}{(a^2 + k_{jk}^2)} d\mathbf{k}_j d\mathbf{k}_k. \tag{2.16}$$

If the quantity

$$\kappa T / \varepsilon_a, \qquad \varepsilon_a = \hbar^2 a^2 / 2m$$

is sufficiently small, we get

$$K_a = (2\pi)^6 l^3 \frac{3\kappa T}{\varepsilon_a} S_a, \qquad S_a = \left(1 - 5\frac{\kappa T}{\varepsilon_a} + 5 \times 7\left(\frac{\kappa T}{\varepsilon^2}\right)^2\right). \tag{2.16'}$$

This is an asymptotic expansion as shown in Appendix I.

(2)   *The evaluation of $\bar{G}_2$.*

The equation (2·10) can be written as

$$\beta^2 J^2 \tilde{G} = \frac{\beta^2 J^2}{2} S_{pq,rs} B_{pq,rs} (G_{pq,ss})^2, \tag{2.17}$$

$$B_{pq,rs} = \frac{e^{-\beta \varepsilon_{rs}} - e^{-\beta \varepsilon_{pq}}}{\beta(\varepsilon_{pq} - \varepsilon_{rs})}, \quad \varepsilon_{xy} = \varepsilon_x + \varepsilon_y. \tag{2.18}$$

By the definition in I–(1·1) we obtain

$$G_{pq,rs} = \theta_{pq,rs} \sum_a \left(\frac{4\pi A}{V}\right) g_a(sq),$$

$$G_{pq,rs} = \theta_{pq,rs} \sum_a \left(\frac{4\pi A}{V}\right) g_a(rq), \tag{2.19}$$

$$g_x(uv) = \frac{1}{x^2 + k_{uv}^2},$$

where $\theta_{pq,rs}$ is the product of four spin functions.   We can carry out the spin summation and obtain

$$\sum_{\text{spin}} (G_{pq,rs})^2 = 2g^2 \sum_a \left(\frac{4\pi A}{V}\right)^2 \left\{ \frac{1}{2}(g_a(sq))^2 + \frac{1}{2}(g_a(rq))^2 + \frac{1}{g} g_a(sq) g_a(rq) \right\} +$$

$$+ 2g^2 \left(\frac{4\pi A}{V}\right)\left(\frac{4\pi B}{V}\right)\left\{ g_a(sq) g_b(sq) + g_a(sq) g_b(rq) \right.$$

$$\left. \mp \frac{1}{g}(g_a(sq) g_b(rq) + g_a(rq) g_b(sq)) \right\} \tag{2.20}$$

$$(-\text{Fermi}, +\text{Bose}).$$

Hence we get

$$\frac{J^2 G_2}{V} = \frac{g^2 J^2}{V} \left[\sum_a \left(\frac{4\pi A}{V}\right)^2 \left(I_{aa}^{(1)} \mp \frac{1}{g} I_{aa}^{(2)}\right) + 2\left(\frac{4\pi A}{V}\right)\left(\frac{4\pi B}{V}\right)(I_{ab}^{(1)} \mp I_{ab}^{(2)}) \right], \tag{2.21}$$

where

$$I_{xy}^{(1)} = S'_{pq,rs} B_{pq,rs} (g_x(sq)g_y(sq)) = S'_{pq,rs} B_{pq,rs} (g_x(rq)g_y(rq)), \quad (2\cdot22)$$

$$I_{xy}^{(2)} = S'_{pq,rs} B_{pq,rs} (g_x(sq)g_y(rq)) = S'_{pq,rs} B_{pq,rs} (g_x(rq)g_y(sq)). \quad (2\cdot23)$$

Replacing the summations with the corresponding integrals, we get

$$I_{aa}^{(1)} = \frac{V^3}{8\pi(4\sqrt{2})} \Lambda^3 \left(\frac{\kappa T}{\varepsilon_a}\right)\left(\frac{\hbar^2}{2m\kappa T}\right)^2 \Psi(a, T),$$

$$\Psi(a, T) = e^{C_a a^2} \int_{C_a}^{\infty} \frac{e^{-x^2}}{x^2} dx, \quad\quad\quad (2\cdot24)$$

$$\varepsilon_a = \frac{\hbar^2}{2m} a^2, \quad C_a = \frac{1}{2}\frac{\varepsilon_a}{\kappa T}$$

and

$$I_{ab}^{(1)} = \frac{V^3}{8\pi} \Lambda^3 \left(\frac{2\pi m\kappa T}{\hbar^2}\right) \frac{1}{a^2-b^2}(\Phi(b, T)-\Psi(a, T)),$$

$$\Phi(a, T) = \frac{1}{a} e^{C_a a^2} \int_{C_a}^{\infty} e^{-y^2} dy. \quad\quad\quad (2\cdot25)$$

The evaluation of $I_{xy}^{(2)}$ is very troublesome, so we approximate them by $I_{xy}^{(1)}$ for the moment.

(3)　*The second virial coefficient.*

Inserting the above results in $(2\cdot8)$ and $(1\cdot14)$. we obtain the following results.

$$B = \mp (N/4\sqrt{2} \, g)\Lambda^{3/2} + B_1, \quad (-\text{Bose}, +\text{Fermi}) \quad (2\cdot26)$$

$$B_1 = -\frac{N}{g^2 \cdot \Lambda^3}\left(-\beta J \frac{G_1}{\Gamma} + \beta^2 J^2 \frac{G^2}{\Gamma} - \cdots\right)$$

$$= \sum_a \frac{N\Gamma_A}{\kappa T}\left\{\frac{1}{2}\left(1 \pm \frac{1}{g}\right) - (\pm)\frac{3}{2g}\frac{\kappa T}{\varepsilon_a}S_a\right\}$$

$$-\left\{\sum_a \frac{N\Gamma_A^2}{\kappa T\varepsilon_a}\frac{a^3}{16\sqrt{2}\pi}\left(\frac{\varepsilon_a}{\kappa T}\right)^{3/2}\Psi(a, T)\right. \quad\quad (2\cdot27)$$

$$\left. -\frac{1}{2}\frac{N\Gamma_A\Gamma_B}{\kappa T\varepsilon_a}\frac{a^4 b^2}{a^2-b^2}(\Phi(b, T)-\Phi(a, T))\right\}\left(1 \pm \frac{1}{g}\right) + \cdots,$$

$$\Gamma_A = 4\pi J\Lambda/a^2. \quad (+\text{Bose}, -\text{Fermi})$$

Our expression for B_1 seems to have very complicated T dependence. But if we factor out $\Gamma_A/\kappa T$, the remaining factor tends to a constant value when T tends to zero. This is easily seen from the incomplete gamma function type character of the functions Ψ, Φ, etc.. This results corresponds to Green's results. Our formula $(2\cdot25)$ is the expansion in powers of J. In order that the second order term is smaller than the first order one, the inequality

$$JA/a^2 < \hbar^2/2m$$

must hold.

If we choose for m the mass of He^4 atom and set $a \sim 10^8$ cm^{-1} rather arbitrarily, we have

$$JA \lesssim 10^{-22} \text{ erg cm.}$$

Unfortunately, we can not compare our results with the experiments easily, because the potential constants JA, a etc. are very sensitive to the method of fitting the potential curve.

§ 3. Conclusion

Because we have stoped with the second order term in J, our formula is insufficient for large interaction energy. But the temperature dependence when T tends to zero can be seen rather clearly. By the present method, the quantum effects can fully be taken into account in principle and the third or higher coefficients may be treated by straightforward calculations. Finally, the author expresses hearty thanks to Dr. K. Hiroike for the kind suggestions* and important criticisms. He kindly communicated his results before publication.

Appendix I. Calculation of (2·16)

We use the following iterated expression for the integrand.

$$\frac{k_{jk}^2}{a^2+k_{jk}^2}=\sum_1^n (-1)^t \frac{k_{jk}^{2t}}{a^{2t}} -(-1)^n \frac{k_{jk}^{2(n+1)}}{(a^2+k_{ik}^2)a^{2n}}. \tag{A1·1}$$

So we get

$$K_a=\sum_1^n(-1)^t \frac{K_{2t}}{a^{2t}}-(-1)^n R_n, \tag{A1·2}$$
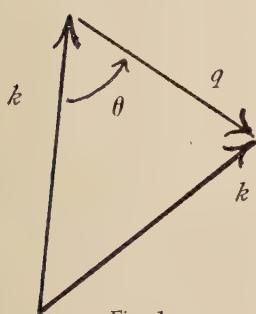
$$K_{2t}=\int d\mathbf{k}e^{-\beta s k^2} \int_0^\infty \int_{-1}^1 2\pi q^{2(t+1)}\exp\left[-\beta s q^2+2\beta k q z\right] dq dz,$$

$$= (2\pi)^6\left(\frac{2\pi mxT}{\hbar^2}\right)^{3+t} \cdot 1 \cdot 3 \cdot 5 \cdots (2t+1), \left(s=\frac{\hbar^2}{2m}\right). \tag{A1·3}$$

Fig. 1.

The relation between integration variables in (A1·3) are shown in Fig. 1. The remainder R_t is always smaller than K_{2t}/a^{2t} in its magnitude and we can take this formula as an asymptotic expansion.

Appendix II. Derivation of (2.24) and (2.25)

(1) *Derivation of* (2·24).

By definition $I_{aa}^{(1)}$ is expressed as

---

* Especially, for the consideration on spin.

$$I_{aa}^{(1)} = \frac{1}{4}\left(\frac{V}{(2\pi)^3}\right)^3 \iiint B_{pq,rs}\,(g_a(sq))^2 dk_p dk_q dk_r, \quad k_r + k_s = k_p + k_q \quad (A2\cdot1)$$

Making use of the integral representention

$$B_{pq,rs} = \int_0^1 e^{-\beta\varepsilon_{pq}} e^{\beta\rho(\varepsilon_{pq}-\varepsilon_{rs})} d\rho \qquad (A2\cdot2)$$

and the integration variables as shown in Fig. 2, we can proceed as follows.

$$I_{aa}^{(1)} = \frac{1}{4}\left(\frac{V}{(2\pi)^3}\right)^3 \int_0^1 d\rho \iiint e^{-\beta\varepsilon_{pq}}$$

$$\cdot\, e^{\beta\rho(\varepsilon_{pq}-\varepsilon_{rs})}(g_a(sq))^2 dk_p dk_q dk_r$$

$$= \frac{1}{4}\left(\frac{V}{(2\pi)^3}\right)^3 2(2\pi)^3 \int_0^1 d\rho \int_0^\infty k_p^2 dk_p$$

$$\cdot\int_0^\infty k_q^2 dk_q \int_0^\infty k^2 dk \int_{-1}^1 dz_1 \int_{-1}^1 dz_2$$

$$\cdot \exp\left[-\omega(k_p^2+k_q^2)\right]\exp\left[-2\omega\rho(k^2-k_p kz_1 - k_q kz_2)\right],$$

where $\omega = \hbar^2/2mxT$, and

$$I_{aa}^{(1)} = \left[\frac{V^3}{2^4\omega^3(2\pi)^5}\right]\int_0^1 d\rho \int_0^\infty \frac{k^2}{(a^2+k^2)^2}\exp\left[-2\omega\rho(1-\rho)k^2\right]dk$$

$$= [\ ,,\ ]\ \sqrt{2\omega}\int_0^\infty x^2 e^{-x^2}dx \int_0^1 d\rho\,\frac{\sqrt{\rho(1-\rho)}}{(2a^2\omega\rho(1-\rho)+x^2)^2}$$

$$= [\ ,,\ ]\ \frac{\pi}{4a^2}\sqrt{\frac{2}{\omega}}\,C_a^2\int_{C_a}^\infty \frac{xe^{-x^2}}{(x^2+C_a^2)^{3/2}} \qquad (A2\cdot4)$$

$$= \frac{V^3}{2^5\sqrt{2}\,\pi}\Lambda^3\left(\frac{\varepsilon_a}{\kappa T}\right)^{1/2}\Psi(a,T),$$

$$\Psi(a,T) = e^{C_a^2}\int_{C_a}^\infty \frac{e^{-x^2}}{x^2}dx,$$

$$\qquad (A2\cdot5)$$

$$C_a = \varepsilon_a/2\kappa T, \quad \varepsilon_a = \hbar^2/2m\kappa T.$$

(2) *Derivation of* (2·25).

From the definition, we have

$$I_{ab}^{(1)} = \frac{1}{4}\left(\frac{V}{(2\pi)^3}\right)^3 \iiint B_{pq,rs}g_a(sq)g_b(sq)dk_p dk_q dk_r \qquad (A2\cdot6)$$

with the same restriction as in $(A2\cdot1)$
Transformimg the integrand as
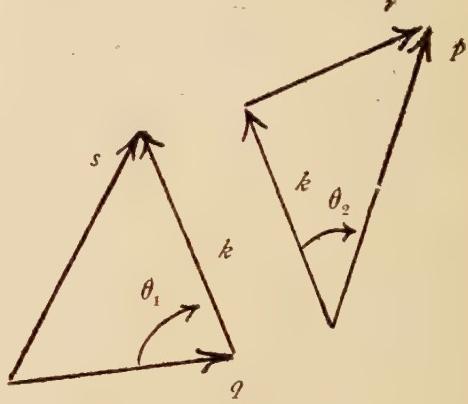
Fig. 2.

$$\frac{1}{(a^2+k_{sq}{}^2)(b^2+k_{sq}{}^2)}=\frac{1}{a^2-b^2}\left(\frac{1}{b^2+k_{sq}{}^2}-\frac{1}{a^2+k_{sq}{}^2}\right)$$

we can go in the similar way as above, and get

$$I_{ab}{}^{(1)}=\frac{1}{4}\left(\frac{V}{(2\pi)^3}\right)^3\frac{(2\pi)^4}{4\omega^4}-\frac{\pi}{a^2-b^2}(\Phi(b,T)-\Phi(a,T)) \qquad (\text{A}2\cdot7)$$

where

$$\Phi(a,T)=\frac{1}{a}e^{c_a{}^2}\int_{c_a}^{\infty}e^{-v^2}dy \qquad (\text{A}2\cdot9)$$

References

1) H. Ichimura, Prog. Theor. Phys. **11** (1954), 374.
2) H. S. Green, Proc. Phys. Soc. **65** (1952), 1022.

528

Statistical Mechanics of Surface Tension of Curved Interface, I

Tomio SEGAWA

Department of Physics, Ehime University, Matsuyama

and   Ei TERAMOTO

Department of Physics, Kyoto University, Kyoto, Japan

(Received November 6, 1953)

From the microscopic point of view, surface tension is defined in two different ways: the statistical-thermodynamic and so to speak, the kinetic. In this paper these two definitions are examined with regard to a cylindrical interface between liquid and vapor, and it is shown that under certain conditions these two definitions become equivalent with each other and also with the usual thermodynamic definition. The methods used are those due to Harasima, MacLellan, and Kirkwood and Buff. The possibility of extending these methods to a spherical interface are briefly mentioned, leaving detailed calculations to be covered by future papers.

§ 1.   Introduction

Surface tension measured as a macroscopic quantity may be defined in two different and distinct ways from the microscopic point of view. The one, statistical-thermodynamic, is defined as:

$$\gamma = \partial F / \partial S_{a\beta} \tag{1}$$

where F is the free energy of the two-phase system considered and $S_{a\beta}$ the area of the interface. Calculations of surface tension based on this concept were made by Fowler, Harasima and MacLellan with regard to plane interfaces. The other definition is, so to speak, kinetic, and is based on calculations of the force across a plane strip perpendicular to the interface. Calculations according to this definition were carried out by Kirkwood and Buff, and MacLellan for the case of a plane interface. These two definitions of surface tension are, of course, mutually independent. But it was shown by Kirkwood and Buff, and MacLellan that in the case of a plane interface these two definitions are equivalent and lead to the same result.

The case of a curved interface presents a number of difficulties. For example, surface tension has two principal values corresponding to the two principal curvatures. Moreover, in the case of a closed surface, it is not clear in the statistical thermodynamic definition which quantities should be kept constant in performing the partial differentiation with respect to the interfacial area. The kinetic definition itself does not present any conceptual difficulty but calculations of stress dyadic for a curved interface become very involved, which makes the method of approximation a subject for discussion.

Among the many curved surfaces the most interesting and important is, of course, the spherical interface. At the present stage however, this form involves too many difficulties to be soluble. For this reason we take as a preliminary step the case of a cylindrical interface, and calculate the surface tension according to the two definitions. The results will then be compared with each other. Quantitative discussions on the curvature dependency of surface tension and the extension of these methods to the spherical interface will be the subjects of future papers.

§ 2.   Surface tension as a thermodynamic quantity

Let us consider a liquid cylinder of radius b and length h, and surrounded by its own vapor. Suppose that the length of the cylinder is increased to $(1+2\varepsilon)h$ and the radius decreased to $(1-\varepsilon)b$. The volume change caused by this deformation is of the order of ε^2. Quantities of this order will henceforth be neglected. The work done during the increase of length may be expressed by

$$2\pi b \cdot 2h\varepsilon\gamma_z,$$

and similarly the work done during the decrease of radius

$$-h(1+2\varepsilon)2\pi b\varepsilon\gamma_\phi.$$

Then the change in the free energy of this liquid cylinder (of the order of ε)

$$\Delta F = 4\pi bh\varepsilon\gamma_z - 2\pi bh\varepsilon\gamma_\phi$$

and the surface tension, measured thermodynamically,

$$\gamma = \Delta F/\Delta S_{\alpha\beta} = 2\gamma_z - \gamma_\phi. \tag{2}$$

Here we have introduced two quantities γ_z and γ_ϕ. These are considered to be the two principal values of the surface tension dyadic, that is, the force across a unit length on a cylindrical surface perpendicular to and parallel to the z-axis respectively. The surface tension measured by macroscopic means is γ in eq. (2), and this can be considered to be composed of two terms γ_z and γ_ϕ from the microscopic standpoint. In fact, as shown in sections III, IV and V, the terms corresponding to the above γ_z and γ_ϕ may be obtained by calculation based on both the statistical thermodynamics and the kinetic theories, when neglecting orders of ε^2 and above.

§ 3.   Surface tension as a statistical thermodynamic quantity

(1)   First we use the method devised by Harasima. For this purpose let us consider a concentric double cylinder, and suppose that the inner cylinder is liquid (α-phase), the outside of which is filled with its own vapor (β-phase). Let $h > b$ so that end effects may be neglected. We deform this concentric double cylinder in the way described at the beginning of section II. Then,

$$\Delta V = O(\varepsilon^2)$$

$$\varDelta S_{\alpha_3}=2\pi b h \mathcal{E}+O\left(\mathcal{E}^2\right) \tag{3}$$

where S_{α_3} is the interfacial area, that is, the lateral surface area of the inner cylinder. We calculate the change of free energy caused by this deformation. The final state is denoted by primes in the following equations:

$$e^{-F/kT}=\frac{1}{N!}\left(\frac{2\pi m k T}{h^2}\right)^{3N/2}\int\cdots\int_{V}e^{-\Phi/kT}\prod_{i=1}^{N}d\boldsymbol{x}^{(i)}, \tag{4}$$

$$e^{-F'/kT}=\frac{1}{N!}\left(\frac{2\pi m k T}{h^2}\right)^{3N/2}\int\cdots\int_{V'}e^{-\Phi'/kT}\prod_{i=1}^{N}d\boldsymbol{x}^{(i)}. \tag{5}$$

To calculate these integrals we use cylindrical coordinates

$$d\boldsymbol{x}=\rho\,d\phi\,d\rho\,dz.$$

Let us introduce new variables θ_1, θ_2 and θ_3 by putting in eq. (4)

$$\phi=2\pi\theta_1,\quad \rho=a\theta_2,\quad z=h\theta_3$$

and in eq. (5)

$$\phi=2\pi\theta_1,\quad \rho=a(1-\mathcal{E})\theta_2,\quad z=h(1+2\mathcal{E})\theta_3.$$

Then eqs. (4) and (5) respectively become

$$e^{-F/kT}=\frac{(2\pi a^2 h)^N}{N!}\left(\frac{2\pi m k T}{h^2}\right)^{3N/2}\int\cdots\int e^{-\Phi/kT}\prod_{i=1}^{N}\theta_2^{(i)}d\theta_1^{(i)}d\theta_2^{(i)}d\theta_3^{(i)} \tag{6}$$

and

$$e^{-F'/kT}=\frac{(2\pi a^2 h)^N}{N!}\left(\frac{2\pi m k T}{h^2}\right)^{3N/2}\int\cdots\int e^{-\Phi'/kT}\prod_{i=1}^{N}\theta_2^{(i)}d\theta_1^{(i)}d\theta_2^{(i)}d\theta_3^{(i)}. \tag{7}$$

Here the domain of integration is a unit cube in the θ-space in both (6) and (7). Moreover, the distance between two molecules i and j can be written in the form

$$r_{ij}=\sqrt{\rho_i^2+\rho_j^2-2\rho_i\rho_j\cos(\phi_j-\phi_i)+(z_j-z_i)^2}, \tag{8}$$

then for the same values of θ_1, θ_2 and θ_3, the corresponding distance at the final state becomes

$$r_{ij}'=r_{ij}+\frac{2z_{ij}^2-\rho_{ij}^2}{r_{ij}}\mathcal{E}+o\left(\mathcal{E}^2\right). \tag{9}$$

And

$$e^{-\Phi'/kT}=e^{-\Phi/kT}\left\{1-\frac{1}{kT}\sum_{i<j}\frac{d\phi_{ij}}{dr_{ij}}\frac{2z_{ij}^2-\rho_{ij}^2}{r_{ij}}\mathcal{E}\right\}. \tag{10}$$

Substituting eq. (10) into eq. (7), we obtain, with the use of the definition of the pair distribution function, $n^{(2)}(\boldsymbol{x}^{(1)},\boldsymbol{x}^{(2)})$,

$$e^{-(F'-F)/kT}=1-\frac{\pi h \mathcal{E}}{kT}\int\cdots\int_{V}\frac{d\phi_{12}}{dr_{12}}\frac{2z_{12}^2-\rho_{12}^2}{r_{12}}n^{(2)}(\rho_1,\boldsymbol{r}_{12})\rho_1\,d\rho_1\,d\boldsymbol{r}_{12}, \tag{11}$$

and finally

$$\gamma = \frac{F' - F}{\varDelta S_{\alpha\beta}} = \frac{1}{2b} \int \cdots \int_V \frac{d\phi_{12}}{dr_{12}} \frac{2z_{12}^2 - \rho_{12}^2}{r_{12}} \, n^{(2)}(\rho_1, \boldsymbol{r}_{12}) \rho_1 \, d\rho_1 \, d\boldsymbol{r}_{12}. \tag{12}$$

Here, in terms of the pair distribution functions $n_\alpha^{(2)}$ and $n_\beta^{(2)}$ in the α- and β-phase, the superficial pair distribution function $n_s^{(2)}(\rho_1, \boldsymbol{r}_{12})$ is introduced as follows:

$$n_{\alpha\beta}^{(2)}(r) = [I - A(\rho)] n_\alpha^{(2)} - A(\rho) n_\beta^{(2)},$$

$$A(\rho) = 0 \text{ for } \rho < b \text{ and } = 1 \text{ for } \rho \geq b,$$

$$n_s^{(2)}(\rho_1, \boldsymbol{r}_{12}) = n^{(2)}(\rho_1, \boldsymbol{r}_{12}) - n_{\alpha\beta}^{(2)}(r).$$

From symmetry considerations eq. (12) becomes

$$\gamma = \frac{1}{2b} \int \cdots \int \frac{d\phi_{12}}{dr_{12}} \frac{2z_{12}^2 - \rho_{12}^2}{r_{12}} [\Gamma_s^{(2)}(\boldsymbol{r}_{12})]_1 \, d\boldsymbol{r}_{12} \tag{13}$$

where

$$[\Gamma_s^{(2)}(\boldsymbol{r}_{12})]_1 = \int_0^\infty n_s^{(2)}(\rho_1, \boldsymbol{r}_{12}) \rho_1 \, d\rho_1. \tag{14}$$

$[\Gamma_s^{(2)}]_1$ corresponds to $[\Gamma_s^{(2)}]_0$ as introduced by Kirkwood and Buff in the case of a plane interface. We must note here the way in which b is chosen. The above consideration does not prescribe b very precisely. This fact is not important in the case of a plane interface but becomes important when considering a curved interface, when for instance the curvature dependency of the surface tension is discussed. This suggests that for the determination of b the above method is not sufficient and that an additional condition should have to be introduced.

(2) Next we use a method devised by MacLellan which is essentially identical with the Harasima's method. However, in MacLellan's paper his equations (17) to (20) appear to the present authors to leave room for some doubt, and these equations cannot be used in their original forms in the case of curved interfaces. Accordingly we will here briefly describe the calculation based on this method and discuss its correspondence to Harasima's method.

We denote the volume of the inner and outer cylinders by V and W respectively,

$$V = \pi b^2 h, \quad W = \pi a^2 h, \quad S_{\alpha\beta} = 2\pi b h. \tag{15}$$

The free energy of this system is given by

$$e^{-F/kT} = \left(\frac{2\pi m k T}{h^2} \right)^{3N/2} Q,$$

$$Q = \frac{1}{N!} \int \cdots \int_W e^{-\Phi/kT} \prod_{i=1}^N dx^{(i)}. \tag{16}$$

Here Q may be considered a function either of a, b and h or of V, W and $S_{\alpha\beta}$. To calculate this integral cylindrical coordinates are again used, but with the unit ρ modified.

$$\phi = 2\pi\theta_1, \quad \rho = b\theta_2, \quad z = h\theta_3.$$

Then,

$$Q = \frac{(2\pi b^2 h)^N}{N!} \int \cdots \int e^{-\Phi/kT} \prod_{i=1}^{N} \theta_2^{(i)} d\theta_1^{(i)} d\theta_2^{(i)} d\theta_3^{(i)}, \qquad (17)$$

$$0 \leq \theta_1 \leq 1, \quad 0 \leq \theta_3 \leq 1, \quad 0 \leq \theta_2 \leq a/b.$$

We now proceed to calculate $\partial Q/\partial b$ and $\partial Q/\partial h$. In comparison with Harasima's method, it is important to note that h and a/b must be kept constant in performing the partial differentiation with respect to b. This means that in differentiating eq. (17) we need not consider the differentiation of the limit of the integral. Thus we obtain

$$\frac{\partial Q}{\partial b} = \frac{2N}{b} Q - \frac{Q}{2bkT} \int_W \int \frac{d\phi_{12}}{dr_{12}} \frac{\rho_{12}^2}{r_{12}} n^{(2)}(\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}) d\boldsymbol{x}^{(1)} d\boldsymbol{x}^{(2)}, \qquad (18)$$

$$\frac{\partial Q}{\partial h} = \frac{N}{h} Q - \frac{Q}{2hkT} \int_W \int \frac{d\phi_{12}}{dr_{12}} \frac{z_{12}^2}{r_{12}} n^{(2)}(\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}) d\boldsymbol{x}^{(1)} d\boldsymbol{x}^{(2)}. \qquad (19)$$

Using the relations

$$\partial F/\partial b = -kT/Q \cdot \partial Q/\partial b, \quad \partial F/\partial h = -kT/Q \cdot \partial Q/\partial h,$$

we find

$$\left(\frac{\partial F}{\partial b} \right)_{h, b/a, T} = -\frac{2NkT}{b} + \frac{1}{2b} S_b,$$

$$\left(\frac{\partial F}{\partial h} \right)_{a,b,T} = -\frac{NkT}{h} + \frac{1}{2h} S_h,$$

$$(20)$$

where

$$S_b = 2\pi h \int \cdots \int \frac{d\phi_{12}}{dr_{12}} \frac{\rho_{12}^2}{r_{12}} [\Gamma_s^{(2)}(\boldsymbol{r}_{12})]_1 d\boldsymbol{r}_{12},$$

$$S_h = 2\pi h \int \cdots \int \frac{d\phi_{12}}{dr_{12}} \frac{z_{12}^2}{r_{12}} [\Gamma_s^{(2)}(\boldsymbol{r}_{12})]_1 d\boldsymbol{r}_{12}.$$

$$(21)$$

With the use of the thermodynamic relations:

$$-P_\alpha = \partial F/\partial V = 1/V \cdot (b \, \partial F/\partial b - h \, \partial F/\partial h),$$

$$\gamma = \partial F/\partial S_{\alpha\beta} = 1/S_{\alpha\beta} \cdot (-b \, \partial F/\partial b + 2h \, \partial F/\partial h),$$

the following expressions for P_α the pressure of the phase α and the surface tension γ are obtained;

$$P_\alpha = NkT/V - 1/2V \cdot (S_b - S_h),$$

$$\gamma = 1/2S_{\alpha\beta} \cdot (2S_h - S_b). \qquad (22)$$

This γ is, with reference to eq. (21), equal to the γ given by Harasima's method, which is as it should be. The advantages of MacLellan's method are : that (a) calculations are

carried out rather mechanically, and (b) physical meanings are given to S_b and S_h thus facilitating concurrence with the phenomenological theory.

§ 4. Surface tension as a kinetic quantity

The method used by Kirkwood and Buff in the case of a plane interface is applied to the same problem as considered above. The principle underlying this method is as follows :

We denote by $\sum'$ the stress dyadic at a point on Gibbs' dividing surface being considered. i, j and k are unit vectors of a right-handed orthogonal coordinate system, and their directions coincide with the principal directsons of $\sum'$. Let i be perpendicular to Gibbs' dividing surface and m a unit vector orthogonal to i. Take a plane strip of unit width perpendicular to m, the two ends of which are in the homogeneous α and β phases respectively. It should then be convenient to employ the term " surface tension across m " for the surface tension obtained from the total force per unit width acting across the plane strip, this definition becomes necessary, from the fact that in general cases the surface tension across m may not be independent of the direction of m. Thus our main problem is to calculate the surface tension with respect to the two principal directions.

In the case of a cylindrical interface, we take as m the axial and lateral directions. First, we take an axial unit vector as m. Then the strip may be replaced by a sector which has an arc of unit length at radius b. γ_z, the surface tension across z, becomes

$$\gamma_z = \sum_{zz} + \frac{1}{b} \int_0^a p(\rho)\, \rho \, d\rho. \tag{23}$$

Utilizing the relations

$$\sum_{zz} = - \frac{kT}{b} \int_0^a n^{(1)}(\rho)\rho d\rho + \sum'_{zz} \tag{24}$$

and

$$d\sum'_{zz} = m \cdot \nabla \phi(r_{12}) n^{(2)}(\rho_1, r_{12}) dr_1 \, dr_{12},$$

$$\sum'_{zz} = \int_0^a \int_0^{1/b} \int_{-\infty}^0 \int_{-z_1}^\infty \int_0^\infty \int_0^\pi \frac{z_{12}}{r_{12}} \frac{d\phi_{12}}{dr_{12}} n^{(2)}(\rho_1, r_{12}) r_{12}\, d\phi_{12}\, d\rho_{12}\, dz_{12}\, \rho_1\, dz_1\, d\phi_1\, d\rho_1,$$

we perform an integration by parts with respect to z and make use of

$$p(\rho) = kTn^{(1)}_{\alpha\beta} - \frac{1}{6} \int r_{12} \frac{d\phi_{12}}{dr_{12}} n^{(2)}_{\alpha\beta}\, dr_{12},$$

and removing the limit $a \to \infty$, we finally obtain

$$\gamma_z = \int_0^\infty \int_0^{1/b} [p - p_{zz}']\rho\, d\phi\, d\rho,$$

$$p_{zz}' = kTn^{(1)}(\rho) - \frac{1}{2} \int \frac{z_{12}^2}{r_{12}} \frac{d\phi_{12}}{dr_{12}} n^{(2)}(\rho_1, r_{12})\, dr_{12}, \tag{25}$$

and

$$\gamma_z = -kT[\Gamma_s^{(1)}]_1 + \frac{1}{2b} \int \frac{z_{12}^2}{r_{12}} \frac{d\phi_{12}}{dr_{12}} [\Gamma_s^{(2)}]_1 \, d\boldsymbol{r}_{12}, \tag{26}$$

where

$$[\Gamma_s^{(1)}]_1 = \frac{1}{b} \int_0^\infty n_s^{(1)}(\rho) \, \rho \, d\rho.$$

$[\Gamma_s^{(1)}]_1$ corresponds to the superficial density in the case of a plane interface. Excepting this term, eq. (26) coincides with a component of γ obtained from statistical thermodynamics.

Next we take as $\boldsymbol{m}$ a unit vector orthogonal to z-axis. (We might call this direction x). This case does not differ in form from the former, but when we integrate $d\sum_{xx}'$ to obtain $\sum_{xx}'$, the integration limit contains integrating variables other than x_1, and integration by parts is not feasible. This difficulty should be overcome, for example, by means of a series expansion of the pair distribution function, but this procedure should be found to be rather tedious. Therefore let us be satisfied with the following simplified consideration. Our purpose is to examine the relationship between the statistical-thermodynamic and the kinetic definitions of the surface tension of a cylindrical interface when considering an approximation of the first order. Accordingly, work done against the stress (that is, the increase of free energy, not the stress itself), must be compared with the surface tension obtained on the basis of statistical thermodynamics. In the present case we must claculate the work done to produce a unit area hypothetically by rotating by an angle $1/b$ a strip of unit width perpendicular to x and passing through the z-axis. As the stress against which the work is done, we take $\sum_{xx}'$, previously obtained for the plane interface. If we denote the stress at ρ by $\sum_{xx}'(\rho)$ the following equation for this work done is derived :

$$\int_0^{1/b} \int_0^a \sum_{xx}'(\rho) \, \rho \, d\rho \, d\phi = \frac{1}{b} \int_0^a \int \frac{x_{12}^2}{r_{12}} \frac{d\phi_{12}}{dr_{12}} n^{(2)}(\rho, \boldsymbol{r}_{12}) \rho d\rho d\boldsymbol{r}_{12}.$$

From symmetry considerations, the same expression is obtained for the y direction. Adding these two expressions and dividing by two, we obtain

$$\int_0^{1/b} \int_0^a \sum_{xx}'(\rho) \rho d\rho d\phi = \frac{1}{2b} \int_0^a \int \frac{\rho_{12}^2}{r_{12}} \frac{d\phi_{12}}{dr_{12}} n^{(2)}(\rho, \boldsymbol{r}_{12}) \rho d\rho d\boldsymbol{r}_{12}.$$

Then, following the same procedure as used to derive (23) and (24) and making use of the equation for $p(\rho)$, we finally have :

$$\gamma_\phi = -kT[\Gamma_s^{(1)}]_1 + \frac{1}{2b} \int \frac{\rho_{12}^2}{r_{12}} \frac{d\phi_{12}}{dr_{12}} [\Gamma_s^{(2)}(\boldsymbol{r}_{12})]_1 d\boldsymbol{r}_{12}. \tag{27}$$

The second term of this equation is the other component of γ obtained from statistical-thermodynamic considerations.

§ 5. Conclusions

We will now discuss the results obtained above. In section III (1), eq. (13) was obtained for surface tension by means of Harasima's method. Though the γ is obviously composed of two terms and is considered correspond to eq. (2); this correspondence is not completely clarified by this method only. However, with the use of MacLellan's method, section III (2), the correspondence

$$\gamma_z \sim S_h \text{ and } \gamma_b \sim S_b$$

becomes clear. Moreover, it is seen from eq. (20) that S_h and S_b are contributions, arising from intermolecular force, to the increase of free energy due to changes in h and b respectively.

In section IV calculations based on the kinetic definition are given, and an expression for γ is obtained. This expression contains a term proportional to $|l'^{(1)}_s|_1$ besides the statistical-thermodynamic term. If we choose b so as to make $[l'^{(1)}_s]_1$ vanish, the kinetic definition leads to the same result as that from the statistical-thermodynamic definition. In addition, the physical meanings of γ_z and γ_b become clear, and correspondence to the phenomenological theory is made straightforward. As to γ_b we shall satisfy ourselves with the first approximation from the plane interface; but the result obtained seems to justify our procedure *post factum*. Throughout the above two treatments the constancy of pressure at points sufficiently distant from the interface is presumed; in other words, b is assumed to be sufficiently large compared to the range of intermolecular force. This assumption is unavoidable so long as statistical considerations are involved. Moreover, in the case of a curved interface a pressure difference exists between the two phases. Although we have not explicitly taken account of this fact, this is postulated in adopting a microsopic standpoint. In the present treatment, howevever, this appears to be concealed behind the approximation used in calculating γ_b.

Finally, we suggest the possibility of extending our considerations to a spherical interface. The methods of Harasima and MacLellan calls for more than two geometrical parameters as independent variables of free energy. For this reason, if we wish to apply this method to the case of a spherical interface, we cannot directly treat a sphere as such, but must first take an ellipsoid or a kind of wedge cut out of the sphere by two or three planes passing through the center. In both cases, however, calculations are complex, and at present we have not succeeded in obtaining a neat expression for the surface tension. Besides, the fact of its being a closed surface presents an additional difficulty and the definition itself of surface tension becomes a point of argument. MacLellan and Kirkwood and Buff recommend the kinetic approach, and we agree with them, but the calculation will not be easily carried out. In this case the approximation which we have used to calculate γ_b in this paper would no longer be valid, and another approximation must be devised.

Acknowledgement

References

1) Fowler, Proc. Roy. Soc. A1£9 (1937), 299.
2) A. Harasima, J. Phys. Soc. Japan **8** (1953), 343.
3) A. G. MacLellan, Proc. Roy. Soc. **A 213** (1952), 274
4) J. G. Kirkwood and F. P. Buff, J. Chem. Phys. **17** (1949), 338.
5) F. P. Buff, J. Chem. Phys. **19** (1951), 1591.

A Formulation of Field Theory in Hilbert Space

Giiti IWATA

Department of Physics, University of Tokyo

(Received December 11, 1953)

A formulation of field theory is set up by making use of tensors in Hilbert space and is applied to quantum electrodynamics. The usual formulation coincides with, as to pair creation, pair annihilation, Compton scattering, and differs from, as to electron self-energy, vacuum polarization, our formulation in which the electron self-energy diverges while the vacuum polarization never appears.

§ 1. Introduction

Quantum theory of fields stands on two feet. One foot is relativity, the other quantum mechanics. Relativity requires the formulation of field theory to be invariant under the Lorentz transformation. As to quantum mechanics, the Hamiltonian formulation due to Heisenberg and Pauli[1] prevails till now either in its original form or in a more refined form. The formulation consists in the application of the quantization method in particle mechanics to the quantization of fields, leaving relativistic features of field theory unstressed. The time there plays a special role, seemingly at variance with the requirement of relativity. The defect has been eliminated by the Tomonaga[2]-Schwinger[3] theory with the inevitable introduction of some subtle and complicated notions such as hypersurface, normal to the surface, condition of integrability.

If one can always pursue each of physical phenomena in its temporal development and if the law of causation is indispensable for the understanding of a succession of physical phenomena, the viewpoint of the Hamiltonian formulation would be rightly considered to be supreme.

It is doubtful however whether one could pursue the behaviour of an elementary particle in space and time so precisely as was described by the Schroedinger equation.

Further the formulation deprives the field theory of its formal beauty which features it before quantization. This is due to the fact that the formulation substitues the Hamiltonian, which is not a relativistic invariant, for the Lagrangian, which is a relativistic invariant, and forsakes the variation principle which makes the classical theory at once clear and simple.

It seems desirable to rebuild the field theory in a clearer form by replacing the Lagrangian in its original place and by reviving the variation principle in its new aspect.

The space-time has long been regarded as a unique frame of reference for the description of physical phenomena, until quantum mechanics introduces a more abstract space, that is, Hilbert space. The time there, however, is an external parameter, so the welding

of relativity and quantum mechanics is not an easy quest.

Therefore it seems preferable to build a Hilbert space fit for relativity and to take the space as a frame of reference.

In Hilbert space there are many representation coordinates related to each other through unitary transformations. We propose here the postulate[1] that the law of physics should be independent of representation coordinates in Hilbert space.

The postulate stands in definite opposition to that of the Hamiltonian formulation, depriving the space-time of its prime importance as the unique frame of reference and eliminating the law of causation as meaningless in elementary processes. Further, the postulate shifts the main task of field theory from pursuing physical phenomena with the lapse of time to finding out each realizable phenomenon as a whole in the entire Hilbert space.

Summing up the above, we present here a scheme to set up a formulation of field theory on three bases, that is, the Lorentz invariance, the representation invariance and the variation principle.

§ 2. Outline

If the scheme is worked out, the field theory assumes an entirely new aspect.

In the first place, we must use tensors in Hilbert space to represent field variables even in the classical theory and set up the Lagrangian L invariant in Hilbert space as well as in Minkowski space. The Lagrangian density integrated over the entire space-time may be taken as the Lagrangian.

In the second place, the variation principle gives field equations in a simple form

$$\delta L = 0 \quad \text{or} \quad \partial L/\partial F \equiv L_F = 0, \quad \partial L/\partial F^\dagger \equiv L_{F\dagger} = 0 , \tag{1}$$

F, $F^\dagger$ being field variables adjoint to each other. Instead of functional derivatives used hitherto in the field theory, here appear partial derivatives alone.

In the third place, invariant commutation relations imposed on field variables yield the general relations

$$[L, F^\dagger] = L_F, \ [L, F] = -L_{F\dagger} . \tag{2}$$

In the fourth place, the variation principle, which asserts the expectation value of the Lagranigian L to be stationary for a realizable state described by Ψ and its adjoint Φ, gives the equations to Ψ, Φ.

$$L\Psi = l\Psi, \quad \Phi L = l\Phi , \tag{3}$$

l being an eigenvalue of L. In other words, a realizable state is conditioned to be an eigenstate of the Lagrangian L.

As a consequence of the relations (2) and the equations (3), there results Ehrenfest's theorem in its field mechanical version

$$\Phi L_F \Psi = 0, \quad \Phi L_{F\dagger} \Psi = 0, \tag{4}$$

which recovers classical field equations (1) in their statistical form.

If our formulation is applied to the quantization of a free field, there results the appearance of virtual quanta. It is disheartening but unavoidable. The virtual quanta, however, contribute nothing to the expectation values of field variables, giving rise to no harmful consequences. If the formulation is modified to exclude virtual quanta, it is difficult to explain scattering processes in which virtual quanta intervene. So they are rather welcome guests than real gate-crashers.

As to quantum electrodynamics, our formulation gives correct matrix elements for pair creation, pair annihilation and Compton scattering. The vacuum polarization never appears in our formulation, owing to the theorem that the number of electron quanta in the entire Hilbert space is invariant throughout every interaction between electron and photon. The self-energy of an electron diverges, however, in our formulation. We fail to eliminate it. It diverges in a different way from the usual one, so that the renormalization technique is not applicable to our case in its usual form. So long as the technique gives not always a definite answer to the problem of field theory, every scheme should not be abandoned for its apparent unfamiliarity.

§ 3. The classical theory in Hilbert space

The Hilbert space used in non-relativistic quantum mechanics is of three parameters. The time there is an external parameter and plays a special role at variance with the requirement of relativity. If we regard the time as another coordinate and introduce Minkowski space, we should build a Hilbert space of four parameters. We take as the four parameters eigenvalues of four commutable operators, for example, x^0, x^1, x^2, x^3 or p_0, p_1, p_2, p_3.

When space-time coordinates x^i are taken as parameters, our Hilbert space is regarded as the space-time. When energy-momentum components p_i are taken, it is regarded as the energy-momentum space. Hilbert space itself is a oneness provided with various representation coordinates.

Transformation functions $(x|p)$, $(p|x)$, which interchange the x-representation with the p-representation, are expressed in the form

$$(x|p) = (2\pi)^{-2} \exp i p_k x^k, \quad (p|x) = (2\pi)^{-2} \exp (-i) p_k x^k.$$

The use of tensors in Hilbert space, however, makes them almost useless.

In the classical field theory, field equations have been derived by the variation principle from the Lagrangian density integrated over some domain in the space-time. The postulate of representation invariance requires the Lagrangian to be independent of representation coordinates, so that the Lagrangian must be a representation-invariant composed of tensors in Hilbert space. In many cases, the usual Lagrangian density integrated over the entire space-time is a representation-invariant, but not always so.

It is to be noted that the differentiation operators used in the usual formulation are replaced by the momentum operators in our formulation, so that the variation principle

gives field equations without the use of partial integration.

An example of Lagrangian is given in § 5 in a previous paper.[5]   Since the differentiation operators are replaced by momentum operators, the variation of the Lagrangian L takes the form

$$\partial L = \sum \mathrm{tr}\ (\partial L/\partial F \cdot \partial F + \partial L/\partial F^\dagger \cdot \partial F^\dagger),$$

$F, F^\dagger$ being a pair of tensors that represent field variables, while the summation ranges over all relevant tensors.   So the variation principle gives field equations in a simple form

$$L_F \equiv \partial L/\partial F = 0 \quad L_{F\dagger} \equiv \partial L/\partial F^\dagger = 0 \qquad\qquad (1)^{bis}$$

without the use of functional derivatives.

If the current j^k is defined by $\partial L/\partial p_k$, it satisfies the conservation equation

$$[p_k,\ j^k] = 0 \qquad\qquad (5)$$

as has been proved by C. Gregory.[6]   When the electromagnetic field is present, the p_k is to be replaced by $\pi_k \equiv p_k - eA_k$, A_k being a component of electromagnetic potentials.   The modified current $j^k = \partial L/\partial \pi_k$ satisfies the modified conservation equation

$$[\pi_k,\ j^k] = 0. \qquad\qquad (5')$$

It is to be noted that while in the absence of electromagnetic field the momentum operator p_k is transformed as $p_k \to S p_k S^{-1}$ under a unitary transformation S, in the presence of electromagnetic field the π_k is assumed to be transformed as $\pi_k \to S\pi_k S^{-1}$, so that the transformation of π_k induces the gauge transformation for electromagnetic potentials.

§ 4.   The spinor and radiation fields

The Lagrangian for spinor and radiation fields is set up in a simpler form

$$\begin{aligned} L &= \phi(\gamma^k \pi_k - m)\psi + 1/4 \cdot \mathrm{tr}\ F_{ik}F^{ik} \\ &= L_M + L_I + L_R \\ L_M &= \phi(\boldsymbol{p} - m)\psi = (\phi|\hat{\xi})(\hat{\xi}|\boldsymbol{p} - m|\hat{\xi}')(\hat{\xi}'|\psi) \\ L_I &= -c\phi\gamma^k A_k\psi = -c\ \mathrm{tr}\ j^k A_k \\ L_R &= 1/4 \cdot \mathrm{tr}\ F_{ik}F^{ik} \end{aligned} \qquad\qquad (6)$$

at the cost of the charge conjugation,[4] where

$$\gamma^i\gamma^k + \gamma^k\gamma^i = 2g^{ik}, \quad g^{00} = 1,$$

$$(\hat{\xi}|j^k|\hat{\xi}') = (\phi|\hat{\xi}')\gamma^k(\hat{\xi}|\psi)$$

$$F_{ik} = \frac{1}{ie}[\pi_i,\ \pi_k] = i[p_i,\ A_k] - i[p_k,\ A_i] - ie[A_i,\ A_k].$$

$$\boldsymbol{p} = \gamma^k p_k.$$

The variation principle gives field equations

$$(\gamma^k \pi_k - m)\psi = 0$$

$$\phi(\gamma^k \pi_k - m) = 0$$

$$i|\pi_i, I^{ki}| - cj^k = 0.$$

If A_k, f_{ik}, j_k commute with each other, we shall have the usual form. Hereafter we assume their commutability without confidence.

Since A_k is hermitean, we decompose it into the sum of two operators a_k, $a_k^\dagger$ hermite-conjugate to each other in order to get the general formula (1). We transform then the Lagrangian L_R, which is

$$L_R = 1/2 \cdot (p - p')^2 (p|A_k|p')(p'|A^k|p) - 1/2 \cdot (p^k - p^{k'})(p_i - p_i')(p|A_k|p')(p'|A^i|p)$$

$$(p - p')^2 = g^{ik}(p_i - p_i')(p_k - p_k')$$

in the p-representation, into

$$L_R = 2(p - p')^2(p|a_k^\dagger|p')(p'|a^k|p) - 2(p^k - p^{k'})(p_i - p_i')(p|a_k^\dagger|p')(p'|a^i|p). \quad (7)$$

We used here the correspondence

$$A^2 = (a + a)^2 \to 4a^\dagger a$$

for which the first reason is to obtain field equations equivalent to old ones, and the second is to get a normal form of the type $a^\dagger a$.

The variation principle gives new field equations

$$2\lfloor p_i \lfloor p^i, a_k \rfloor \rfloor - 2\lfloor p_k \lfloor p^i, a_i \rfloor \rfloor - cj_k = 0$$

$$2\lfloor p_i \lfloor p^i, a_k^\dagger \rfloor \rfloor - 2\lfloor p_k \lfloor p^i, a_i^\dagger \rfloor \rfloor - cj_k = 0$$

which, combined together, give evidently old equations

$$[p_i [p^i, A_k]] - [p_k[p^i, A_i]] - cj_k = 0 .$$

If the Lorentz condition $[p^i, A_i] = 0$ or $[p^i, a_i] = |p^i, a_i^\dagger| = 0$ is imposed, these equations get simpler.

§ 5. Remarks on the non-local theory

There is a scalar field represented by mixed tensors of rank 2, $(\bar{\xi}|U|\bar{\xi}')$, $(\bar{\xi}|U^\dagger|\bar{\xi}')$, while the scalar field in the previous paper[5] is represented by vectors.

The Lagrangian of the scalar field of rank 2 is given by a Lorentz and representation invariant

$$L = (\bar{\xi}|[U^\dagger, p_k][p^k, U] - m^2 U^\dagger U|\bar{\xi})$$

$$= \text{tr} \{[U^\dagger, p_k][p^k, U] - m^2 U^\dagger U\},$$

which takes a form

$$L = \{(p - p')^2 - m^2\} (p|U^\dagger|p')(p'|U|p)$$

in the p-representation.

The scalar field introduced above is similar to the scalar field that has been given by H. Yukawa,[7] who wished to introduce the size of elementary particle in some sense. Two suffices $\hat{s}$, $\hat{s}'$ of $(\hat{s}|U|\hat{s}')$, however, denote merely the transformation property under representation transformations and have no relation to the size of elementary particle in our scheme. Every observable in quantum mechanics has two suffices. The momentum of a particle is an observable and has two suffices. The electromagnetic potential, which should have the same transformation property with that of the momentum, is also an observable and has two suffices. But it seems curious to think of the size of photon.

If we attempt to reformulate in tensor form the field equations of non-local theory proposed by european physicists P. Kristensen and C. Møller,[8] we meet a difficulty.

The field equations run in their notation

$$(\gamma_\mu \partial/\partial x'^\mu + M)\psi(x') = -g\int F(x', x'', x''')u(x'')\psi(x''')\,dx''dx'''$$

$$(\square'' - m^2)u(x'') = g\int F(x', x'', x''')\bar{\psi}(x')\psi(x''')\,dx'dx'''.$$

We use the abbreviation

$$i\gamma_\mu p_\mu + M = R, \quad -p_\mu p_\mu - m^2 = Q$$

and regard $\psi(x) \equiv (x|\psi)$, $\bar{\psi}(x) \equiv (\psi^\dagger|x)$ as vectors in Hilbert space with confidence. If we denote the number of contravariant suffices minus that of covariant suffices of a tensor A by $\sigma(A)$, we see

$$\sigma(R) = 0, \quad \sigma(Q) = 0, \quad \sigma(\psi) = 1, \quad \sigma(\psi^\dagger) = -1.$$

The σ-values of both members of an equation must be equal to each other, so we get the conditions

$$0 + 1 = \sigma(F) + \sigma(u) + 1$$

$$0 + \sigma(u) = \sigma(F) - 1 + 1.$$

Therefore we have the vanishing of both $\sigma(F)$ and $\sigma(u)$. Since both u and F have odd numbers of suffices, both $\sigma(u)$ and $\sigma(F)$ never vanish. This is the difficulty mentioned above. An evasion out of this difficulty is to assume F, u to be of rank 4, 2 respectively. Instead of the equations quoted above, we shall have then

$$(x|R|x')\,(x'|\psi) = -g\,(xx'|F|x''x''')\,(x''|u|x')\,(x'''|\psi)$$

$$(x'||p_\mu|p_\mu, u||-m^2u|x'') = g\,(\psi^\dagger|x)\,(xx'|F|x''x''')\,(x'''|\psi)$$

which are deducible from the Lagrangian

$$L = (\psi^\dagger|x)\,(x|R|x')\,(x'|\psi) - 1/2 \cdot (x|[u, p_\mu][p_\mu, u] - m^2u^\circ|x)$$

$$+ g(\psi^\dagger|x)\,(xx'|F|x''x''')\,(x''|u|x')\,(x'''|\psi).$$

The use of tensors in Hilbert space gives some restrictions on the non-local theory.

§ 6. The quantization in phase space

So far the use of tensors in Hilbert space does not seem to cause a serious disaccord

between the usual formulation of field theory and ours.

As to the quantization procedure, however, our formulation stands in definite opposition to the Hamiltonian formulation as will be seen later.

In a previous paper[5] classical particle mechanics was shown to be the representation of the contact transformation. The transition from classical mechanics to quantum mechanics was made there by establishing commutation relations invariant under the unitary transformation that is contained in the contact transformation. Commutation relations may be established more definitely by requiring their invariance under the symplectic transformation.

The equation of motion in particle mechanics can be cast in the form

$$\delta\omega_d - d\omega_\delta = 0$$

$$\omega_d = \sum_{i=1}^{n} p_i \, dq^i - H \, dt.$$

If we replace the minus Hamiltonian $-H$ by the momentum p_0 conjugate to the time $t = q^0$ and set up the equation

$$F(q, p) = p_0 + H(q, p) = 0,$$

we get the equation of motion in a homogeneous form

$$\delta\omega_d - d\omega_\delta = \sum_{r=0}^{n}(dq^r \, \delta p_r - dp_r \, \delta q^r) = 0$$

with a subsidiary condition

$$F = 0 ,$$

so that (dq^r, dp_r) can be regarded as a contravariant vector of the contact transformation. We confine the contact transformation group to its linear subgroup, that is, the symplectic group. Then we can regard $g^r = x^r$, $p_r = x^{r\prime}$ as components of a contravariant vector $(x^r, x^{r\prime})$ of the symplectic group.

We denote the skew product[6] of two contravariant vectors $(x^r, x^{r\prime})$, $(y^r, y^{r\prime})$ by

$$\sum_{r}(x^r y^{r\prime} - x^{r\prime} y^r) \equiv J_{\alpha\beta} \, x^\alpha y^\beta, \quad J_{\alpha\beta} = -J_{\beta\alpha}$$

α, β running over $0, 1, \cdots, n, 0', 1', \cdots, n'$ and define covariant components of $(y^r, y^{r\prime})$ by

$$y_\alpha = J_{\alpha\beta} \, y^\beta, \quad \text{or} \quad y_r = y^{r\prime}, \quad y_r' = -y^r.$$

Further we define $J^{\alpha\beta}$ by

$$J_{\alpha\beta} \, x^\alpha y^\beta = J^{\alpha\beta} \, x_\alpha y_\beta$$

or $\qquad J_{\alpha\beta} = J^{\alpha\beta}.$

There is no contravarinat tensor of rank 2 that retains its value under the symplectic transformation, other than a multiple of $J^{\alpha\beta}$.

We regard now x^α as non-commutable operators and establish commutation relations among them.

A contravariant tensor of rank 2 constructed with x^α is of a type

$$bx^\alpha x^\beta + cx^\beta x^\alpha = F^{\alpha\beta}$$

b, c being constants. We require the tensor $F^{\alpha\beta}$ to retain its value under the symplectic transformation. $F^{\alpha\beta}$ must be then a multiple of $J^{\alpha\beta}$. As $J^{\alpha\beta}$ is alternate in indices α, β, $F^{\alpha\beta}$ is also alternate. So c must be equal to $-b$. Therefore we have commutation relations

$$x^\alpha x^\beta - x^\beta x^\alpha = a J^{\alpha\beta} \tag{8}$$

which coincide with usual ones, a being a constant.

Substituting q^r for x^r, p_r for $x^{r\prime}$ we get

$$q^r q^s - q^s q^r = 0, \quad p_r p_s - p_s p_r = 0$$
$$q^r p_s - p_s q^r = a\partial_s^r$$

where a is to be put i, $\partial_s^r = 1$ for $r = s$, $\partial_s^r = 0$ for $r \neq s$. The Schroedinger equation runs

$$F(q, p)\psi = \{p_0 + H(q, p)\}\psi = 0.$$

From the earlier stage of quantum mechanics, the contact transformation has been compared to the unitary transformation. The two transformations are of different natures. While the contact transformation in classical mechanics is of finite dimensions, the unitary transformation in quantum mechanics is of infinite dimensions. While the contact transformation gives commutation relations as was shown above, the unitary transformation leaves any functional relations among observables invariant, so that it gives no restriction on commutation relations.

The symplectic and orthogonal groups are twins having the unitary group as their intersection.[9]

The orthogonal group can be defined as such that leaves the scalar product

$$\sum(x^r y^{r\prime} + x^r y^r) \equiv H_{\alpha\beta} \, x^\alpha \, y^\beta, \quad H_{\alpha\beta} = H_{\beta\alpha}$$

invariant. Commutation relations invariant under the orthogonal transformation are shown to be

$$x^\alpha \, x^\beta + x^\beta \, x^\alpha = a H^{\alpha\beta} \tag{9}$$

along the same line of reasoning as above, where $H^{\alpha\beta} = H_{\alpha\beta}$.

The plus or minus sign that appears in commutation relations is in close relation to the group that underlies them.

Commutation relations (8), (9) are expressed in tensor form of the unitary transformation as

$$
\begin{aligned}
x^r x^s - \rho x^s x^r &= 0 & \rho = 1 \text{ for } (8) \\
x^{r\prime} x^{s\prime} - \rho x^{s\prime} x^{r\prime} &= 0 & \rho = -1 \text{ for } (9) \\
x^r x^{s\prime} - \rho x^{s\prime} x^r &= a\partial_s^r
\end{aligned}
\right\} \tag{10}
$$

where the x^r and $x^{r\prime}$ are regarded as contravariant and covariant vectors of the unitary transformation that leaves $\sum x^r x^{r\prime}$ invariant.

§ 7. The quantization in Hilbert space

In order to establish commutation relations for tensors in Hilbert space, we follow a parallel line of reasoning to that of the preceding section. The unitary transformation in phase space is now to be replaced by the unitary transformation in Hilbert space.

The correspondence

$$x^\nu \to (\xi|F), \quad x^{\nu\prime} \to (F^\dagger|\xi)$$

gives commutation relations

$$\left.\begin{aligned}
(\xi|F)(F^\dagger|\xi') - \rho(F^\dagger|\xi')(\xi|F) &= (\xi|\xi') \\
(\xi|F)(\xi'|F) - \rho(\xi'|F)(\xi|F) &= 0 \\
(F^\dagger|\xi)(F^\dagger|\xi') - \rho(F^\dagger|\xi')(F^\dagger|\xi) &= 0 ,
\end{aligned}\right\} \qquad (11)$$

where the case $\rho = 1$ applies to Boson fields, the case $\rho = -1$ to Fermion fields. The constant a that appears in (10) is put 1.

For tensors of rank 2, we have

$$\left.\begin{aligned}
(\xi|F|\xi')(\xi''|F^\dagger|\xi''') - \rho(\xi''|F^\dagger|\xi''')(\xi|F|\xi') &= (\xi|\xi''')(\xi''|\xi') \\
(\xi|F|\xi')(\xi''|F|\xi''') - \rho(\xi''|F|\xi''')(\xi|F|\xi') &= 0 \\
(\xi|F^\dagger|\xi')(\xi''|F^\dagger|\xi''') - \rho(\xi''|F^\dagger|\xi''')(\xi|F^\dagger|\xi') &= 0 .
\end{aligned}\right\} \qquad (12)$$

For a Boson field, commutation relations and the equality

$$[A_1 A_2 \cdots A_n, B] = [A_1, B]A_2 \cdots A_n + A_1[A_2, B]\cdots A_n + \cdots + A_1 \cdots A_{n-1}[A_n, B]$$

gives the relation

$$[(\xi'|F)(\xi''|F)\cdots(\xi|F^n), (F^\dagger|\xi)] = (\xi'|\xi)(\xi''|F)\cdots(\xi^n|F) + (\xi'|F)(\xi''|\xi)\cdots(\xi^n|F) + \cdots + (\xi'|F)\cdots(\xi^{n-1}|F)(\xi^n|\xi). \qquad (13)$$

Hence to make the commutator of a quantity with $(F^\dagger|\xi)$ is equivalent to differentiate the quantity with respect to $(\xi|F)$

$$\left.\begin{aligned}
[L, (F^\dagger|\xi)] &= L_{\cdot(\xi|F)} \\
{} \\
[L, (\xi|F)] &= -L_{(F^\dagger|\xi)} .
\end{aligned}\right\} \qquad (2')$$

and quite similarly

For a Fermion field, we have at first the relation

$$[(F^\dagger|\xi')(\xi''|F), (F^\dagger|\xi)] = (F^\dagger|\xi')(\xi''|\xi)$$

$$[(F^\dagger|\xi')(\xi''|F), (\xi|F)] = -(\xi|\xi')(\xi''|F)$$

which, combined with the above equality, gives the relation (2') if the L is multi-bilinear in pairs of $F^\dagger$, F.

When field variables $F^\dagger$, F have two or more suffices, we have the same relation

$$[L, F^\dagger] = L_F, \quad [L, F] = -L_{F^\dagger} \qquad (2^{bis})$$

along the same line of reasoning as above.

After quantization is over, the restriction to be imposed on a realizable state should be given.

The variation principle requires the expectation value of the Lagrangian L to be stationary for a realizable state described by Ψ and its adjoint Φ, that is,

$$\delta(\Phi L \Psi / \Phi \Psi) = 0 .$$

Therefore we have the equations to Ψ, Φ

$$L\Psi = l\Psi, \quad \Phi L = l\Phi , \tag{3^{bis}}$$

l being an eigenvalue of the operator L. Hence we can say that a realizable state is an eigenstate of L.

In particle mechanics there exists Ehrenfest's theorem[10] which states that the expectation values of dynamical variables satisfy classical equations of motion.

The relation (2bis) and the equations (3bis) combine to give Ehrenfest's theorem in its field mechanical version. In fact, the expectation value of L_F vanishes for an eigenstate of L, since

$$
\left.
\begin{aligned}
\Phi L_F \Psi &= \Phi[L, F^\dagger]\Psi \\
&= \Phi[l, F^\dagger]\Psi = 0,
\end{aligned}
\right\} \tag{4^{bis}}
$$

and quite similarly

$$\Phi L_{F^\dagger}\Psi = 0 .$$

Therefore classical field equations reappear in their statistical form, seeing that they are derived from the Lagrangian L by the variation principle to take the form

$$L_F = 0, \quad L_{F^\dagger} = 0 \tag{1^{bis}}$$

without the use of functional derivatives.

The conservation equation of the current, which holds in the classical theory, holds also in the quantized theory in its statistical form. Since the quantized Lagrangian also remains unaltered under the unitary transformation, there exists an equality. If the variations due to an infinitesimal unitary transformation are marked by Δ, the variation of the Lagrangian L, which is to vanish, takes the form

$$\Delta L = \mathrm{tr} \left\{ \sum | - \Delta F^\dagger | L, F | + | L, F^\dagger | \Delta F | + \Delta \pi_k\, j^k \right\} .$$

If the infinitesimal transformation is put $S = 1 + H$, H: antihermitian, ΔL reduces to

$$\Delta L = \mathrm{tr}\ H\{[L, Y] + [\pi_k, j^k]\}$$

so that we get the equality

$$| L, Y | + | \pi_k, j^k | = 0$$

where Y is a certain mixed tensor of rank 2 composed of field variables. For example, when F, $F^\dagger$, are contravariant and covariant vectors, we see

$$(\tilde{\xi}|\varDelta F) = (\tilde{\xi}|H|\tilde{\xi}')(\tilde{\xi}'|F), \quad (\varDelta F^{\dagger}|\tilde{\xi}) = -(F^{\dagger}|\tilde{\xi}')(\tilde{\xi}'|H|\tilde{\xi})$$

$$\mathbf{tr}\{-\varDelta F^{\dagger}[L, F]+[L, F^{\dagger}]\varDelta F\} = (\tilde{\xi}'|H|\tilde{\xi})[L, (F^{\dagger}|\tilde{\xi}')(\tilde{\xi}|F)]$$

$$\therefore \quad (\tilde{\xi}|Y|\tilde{\xi}') = (F^{\dagger}|\tilde{\xi}')(\tilde{\xi}|F).$$

When $F, F^{\dagger}$ are mixed tensors of rank 2, we have

$$(\tilde{\xi}|Y|\tilde{\xi}') = (\tilde{\xi}''|F^{\dagger}|\tilde{\xi}')(\tilde{\xi}|F|\tilde{\xi}'') - (\tilde{\xi}|F^{\dagger}|\tilde{\xi}'')(\tilde{\xi}''|F|\tilde{\xi}').$$

The above equality gives the conservation equation in its statistical form since

$$\varPhi[\pi_k, j^k]\varPsi = -\varPhi[L, Y]\varPsi = -\varPhi[l, Y]\varPsi = 0 \tag{5''}$$

for any eigenstate of L.

§ 8. The quantization of the scalar field

The Lagrangian of the scalar field[5] becomes in the p-representation (with a slight change in notation)

$$L = (U^{\dagger}|p)(p|Q|p')(p'|U)$$
$$= Q(p)(U^{\dagger}|p)(p|U) \quad Q(p) = \boldsymbol{p}^2 - m^2 = p^2 - m^2.$$

We impose on $(U^{\dagger}|p)$, $(p|U)$ commutation relations (11), which take the following form in the p-representation

$$\left.\begin{aligned} (p|U)(U^{\dagger}|p') - \rho(U^{\dagger}|p')(p|U) &= (p|p') \\ (p|U)(p'|U) - \rho(p'|U)(p|U) &= 0 \\ (U^{\dagger}|p)(U^{\dagger}|p') - \rho(U^{\dagger}|p')(U^{\dagger}|p) &= 0. \end{aligned}\right\} \tag{11'}$$

Since any two terms $(U^{\dagger}|p)(p|U)$, $(U^{\dagger}|p')(p'|U)$ with fixed values of p, p' commute with each other, the L is diagonal.

Eigenstates of L are found to be

$$\varPsi = (U^{\dagger}|p)^n \cdot 1$$

$$l = nQ(p), \qquad n = 0, 1 \qquad \text{for } \rho = -1$$
$$= 0, 1, 2, \cdots \quad \text{for } \rho = 1$$

for every p-value, except for normalization factors.

The eigenstate $\varPsi = 1$ represents the state where no quantum exists.

While a free quantum has the momentum satisfying the condition $Q(p) = 0$, a quantum which appears in intermediate states of the usual formulation has the momentum not satisfying $Q(p) = 0$. If a quantum $Q(p) \gtrless 0$ is called a virtual quantum, our formulation makes appear a vast multitude of virtual quanta other than a multitude of real quanta $Q(p) = 0$. The consequence is disheartening but unavoidable. We should be forced to consider our formulation purely formal, misguided and misleading, were it not for Ehrenfest's theorem. The postulate of relativity of representation coordinates might turn out to be

nothing but a mathematical fairy tale concocted by a mere speculation, if our formulation should give a result wholly inconsistent with the experiment.

The expectation values of field variables satisfy, however, classical field equations by virtue of Ehrenfest's theorem, so we have

$$\Phi\lfloor L, \ (U^\dagger|p) \rfloor \Psi = \Phi(p|Q|p')(p'|U)\Psi$$
$$= Q(p)\,\Phi(p|U)\,\Psi = 0$$

for an eigenstate Ψ, Φ. Therefore the expectation value of field variable $(p|U)$ vanishes unless the condition $Q(p)=0$ is satisfied, so that the appearance of virtual quanta does not seem absolutely prohibitive.

§ 9. The quantization of spinor field

Commutation relations invariant under the reflections, the spin transformation and the representation transformation are written in the p-representation

$$
\left.
\begin{aligned}
(p|\psi_\alpha)(\phi_\beta|p') + (\phi_\beta|p')(p|\psi_\alpha) &= \{(p|\psi_\alpha),(\phi_\beta|p')\} = \delta_{\alpha\beta}\cdot(p|p') \\
(p|\psi_\alpha)(p'|\psi_\beta) + (p'|\psi_\beta)(p|\psi_\alpha) &= \{(p|\psi_\alpha),(p'|\psi_\beta)\} = 0 \\
(\phi_\alpha|p)(\phi_\beta|p') + (\phi_\beta|p')(\phi_\alpha|p) &= \{(\phi_\alpha|p),(\phi_\beta|p')\} = 0,
\end{aligned}
\right\} \quad (14)
$$

since there is no invariant spinor of rank 2 except for the unit spinor $\delta_{\alpha\beta}$ under reflections and the spin transformation.

To bring the Lagrangian in diagonal form, which is not diagonal even in the p-representation, we introduce the matrix K of four dimensions and its inverse, defined by

$$K = (\boldsymbol{p}\gamma^0 + \mu)\,(2\mu^2 + 2\mu p_0)^{-1/2}$$
$$K^{-1} = (\gamma^0\boldsymbol{p} + \mu)\,(2\mu^2 + 2\mu p_0)^{-1/2}$$
$$\mu(p)^2 = \boldsymbol{p}^2 = p_0^2 - p_1^2 - p_2^2 - p_3^2,$$

where the matrix γ^0 is taken to be diagonal.
The matrix K transforms the momentum $\boldsymbol{p}$ in diagonal form

$$K^{-1}\boldsymbol{p}K = \mu\gamma^0.$$

The change of field variables

$$
\left.
\begin{aligned}
(p|\psi) &= K(p)(p|u), & (\phi|p) &= (v|p)K^{-1}(p) \\
(p|\psi_\alpha) &= K_{\alpha\beta}(p)(p|u_\beta), & (\phi_\rho|p) &= (v_\sigma|p)K^{-1}_{\sigma\rho}(p)
\end{aligned}
\right\} \quad (15)
$$

or

makes the Lagrangian L_M (8) diagonal

$$
\left.
\begin{aligned}
L_M &= (v|p)\{\mu(p)\gamma^0 - m\}(p|u) \\
&= q_\alpha(p)(v_\alpha|p)(p|u_\alpha) \\
&\quad q_\alpha(p): \text{ eigenvalue of } \boldsymbol{p}-m
\end{aligned}
\right\} \quad (16)
$$

leaving the commutation relations (14) invariant,

$$\{(p|u_\alpha),(v_\beta|p')\}=\delta_{\alpha_3}\cdot(p|p')$$
$$\{(p|u_\alpha),\ (p'|u_3)\}=\{(v_\alpha|p),(v_\beta|p')\}=0.$$

(17)

The $(v_\alpha|p)(p|u_\alpha)$ with α, p fixed represents the number of spinor quanta of momentum p in the state α. If we introduce the density ρ by $(\xi|\rho|\xi')=(\phi|\xi')(\xi|\phi)$, the trace of which, that is, $N=\mathrm{tr}\ \rho=(v_\alpha|p)(p|u_\alpha)$, represents the total number of spinor quanta in the entire Hilbert space.

If the diagonal elements of the matrix γ^0 are denoted by $\mathcal{E}_\alpha$, $\mathcal{E}_1=\mathcal{E}_2=-\mathcal{E}_3=-\mathcal{E}_4=1$, the eigenvalues $q_\alpha(p)$ are

$$q_\alpha(p)=\mu(p)\mathcal{E}_\alpha-m\ .$$

Since $\mu(p)^2=\boldsymbol{p}^2=p_0^2-p_1^2-p_2^2-p_3^2$ assumes any real values, $\mu(p)$ will be real as well as imaginary so that the convention $\mu(p)=+(\boldsymbol{p}^2)^{1/2}$ loses its validity. Indeed, $\mu(p)$ is a two-valued function of p_k having the light cone $\boldsymbol{p}^2=0$ as its branch surface. When $p_1=p_2=p_3=0$, $\mu(p)$ is equal to p_0.

Hitherto a quantum of positive energy $p_0>0$ has been called electron, one of negative energy positron. Neither the number of electrons nor that of positrons is a representation-invariant notion. From the Hamiltonian point of view, the distinction between electron and positron is necessary. From our point of view, the distinction is not only clumsy but useless. Hereafter we regard a positron as an electron of negative energy. So the energy of an electron may be positive as well as negative in the following.

A quantum described by $(v_\alpha|p)$ is called real when $q_\alpha(p)=0$, virtual when $q_\alpha(p)\neq0$.

The total number N denotes the total number of electrons, real as well as virtual, irrespective of spin-value.

If we preserve field variables referring to real quanta discarding those referring to virtual ones, the Lagrangian will be reduced to nothing because of the vanishing of $q_\alpha(p)$ for real quanta.

§ 10. The quantization of radiation field

Commutation relations to be imposed on $a_k^\dagger$, a^i are

$$[(p|a^i|p'),(p''|a^k|p''')]=[(p|a_i^\dagger|p'),(p''|a_k^\dagger|p''')]=0$$
$$[(p|a^i|p'),\ (p''|a_k^\dagger|p''')]=\delta_k^i(p|p''')(p''|p').$$

(18)

The Lagrangian L_R (7) with these commutation relations is not diagonal, so we introduce 4 mutually orthogonal unit vectors labeled by the greek suffix λ, ranging over 0, 1, 2, 3 among which the vector labeled by $\lambda=0$ is taken to be parallel to the momentum $p-p'$. If we denote components of 4 unit vectors by

$$S_\lambda^i(p,p'),$$

in particular

$$S_0^i=(p^i-p^{i'})/\mu(p-p'),\ S_i^0=(p_i-p_i')/\mu(p-p')$$

(19)

we have

$$S^i_\lambda S_i^\mu = \delta^\mu_\lambda. \quad [S_i^\mu = g_{ik} g^{\mu\lambda} S^k_\lambda].$$

Making use of the 4 unit vectors as a local coordinate system, we decompose the potentials a^i, $a_k^\dagger$ as

$$\begin{aligned}(p'|a^i|p) &= S^i_\lambda(p, p')(p'|c^\lambda|p) \\ (p|a_k^\dagger|p') &= S_k^\mu(p,p')(p|c_\mu^\dagger|p')\end{aligned} \quad \Bigg\} \tag{20}$$

getting

$$L_R = 2(p-p')^2 S_k^\mu(p|c_\mu^\dagger|p') S^k_\lambda(p'|c^\lambda|p) - 2(p^k-p^{k'})(p_i-p_i') S_k^\mu(p|c_\mu^\dagger|p') S^i_\lambda(p'|c^\lambda|p)$$

$$= 2(p-p')^2(p|c_\mu^\dagger|p')(p'|c^\mu|p) - 2(p-p')^2(p|c_0^\dagger|p')(p'|c^0|p)$$

$$= 2(p-p')^2 \sum_{\mu=1}^{2}(p|c_\mu^\dagger|p')(p'|c^\mu|p). \tag{21}$$

Thus the term labeled by $\mu = 0$ disappears out of the Lagrangian, as a consequence of gauge invariance. This is the elimination of longitudinal photon. The commutation relations for c^λ, $c_\mu^\dagger$ are of the same form as (18), so that the Lagrangian is now diagonal. For fixed values of p, p', λ, the $(p|c_\lambda^\dagger|p')(p'|c^\lambda|p)$ represents the number of λ-polarized photons of momentum $p-p'$.

Because the 4 unit vectors are not uniquely determined, an arbitrariness might appear. So we use, instead of c^λ, $c_\mu^\dagger$, modified potentials b^k, $b_k^\dagger$ that are obtained from a^k, $a_k^\dagger$ by dropping off their respective c^0- or $c_0^\dagger$-components so as to satisfy the Lorentz condition. Indeed, modified potentials b^k defined by

$$(p'|b^k|p) = (p'|a^k|p) - S^k_0(p, p')(p'|c^0|p)$$

$$= (p'|a^k|p) - S^k_0(p, p') S_i^0(p, p')(p'|a^i|p)$$

$$= (p'|a^k|p) - \frac{(p^k-p^{k'})(p_i-p_i')}{(p-p')^2}(p'|a^i|p)$$

satisfy the Lorentz condition

$$[p_k, b^k] = 0. \tag{22}$$

The $b_k^\dagger$ defined by

$$(p|b_k^\dagger|p') = (p|a_k^\dagger|p') - \frac{(p_k-p_k')(p^i-p^{i'})}{(p-p')^2}(p|a_i^\dagger|p')$$

satisfy also the Lorentz condition $[p^k, b_k^\dagger] = 0$.
Commutation relations for $b_k^\dagger$, b^i become

$$\begin{aligned}[(p|b^i|p'),(p''|b^k|p''')] &= [(p|b_i^\dagger|p'),(p''|b_k^\dagger|p''')] = 0 \\ [(p|b^i|p'),(p''|b_k^\dagger|p''')] &= (\delta^i_k - \frac{(p^i-p^{i''})(p_k-p_k')}{(p-p')^2})(p|p''')(p''|p'),\end{aligned} \quad \Bigg\} \tag{23}$$

while the Lagrangian takes the form

$$L_R = 2(p-p')^2(p|b_k^\dagger|p')(p'|b^k|p). \tag{24}$$

§ 11. Quantum electrodynamics

The total Lagrangian of electron and radiation fields is now expressed in the form

$$L = L_M + L_R + L_I$$
$$L_M = q_\alpha(p)\,(v_\alpha|p)\,(p|u_\alpha)$$
$$L_R = 2\,(p-p')^2\,(p|b_k^\dagger|p')\,(p'|b^k|p) \tag{25}$$
$$L_I = -c\,(p|j^k|p')\,(p'|A_k|p) = -c\,(v|p')\,a^k(p',p)\,(p|u)\,(p'|A_k|p)$$

where $a^k(p',p) = K^{-1}(p')\gamma^k K(p)$, or $a^k_{\alpha\beta}(p',p) = K^{-1}_{\alpha\rho}(p')\gamma^k_{\rho\sigma}K_{\sigma\beta}(p)$

We modify L_I as follows

$$L_I = -c\,(p|j^k|p')\,(p'|a_k + a_k^\dagger|p)$$
$$= -c\,(p|j^k|p')\left\{(p'|b_k + b_k^\dagger|p) + \frac{(p_k - p_k')(p^i - p^{i'})}{(p-p')^2}(p'|A_k|p)\right\}$$
$$= -c\,(p|j^k|p')\,(p'|b_k + b_k^\dagger|p) + c\,\frac{1}{(p-p')^2}(p|[p_k, j^k]|p')\,(p'|[p^i, A_i]|p).$$

In the classical theory, the last term disappears as a result of the conservation of current if $[A_k, j^k] = 0$.

If we use the relations (19), (20), we have

$$(p'|c^0|p) = S_i^0(p, p')\,(p'|a^i|p) = -\frac{1}{\mu(p-p')}(p'|[p_i, a^i]|p)$$

$$(p|c_0^\dagger|p') = S^k_0(p|a_k^\dagger|p') = \frac{1}{\mu(p-p')}(p|[p^k, a_k^\dagger]|p'),$$

so that $$(p|c_0^\dagger - c^0|p') = \frac{1}{\mu(p-p')}(p|[p^k, A_k]|p').$$

Since the $c_0^\dagger$, c^0 never appear elsewhere in the total Lagrangian (25) and any two terms $(p|c_0^\dagger - c^0|p')$, $(p''|c_0^\dagger - c^0|p''')$ commute with each other, the term $(p|c_0^\dagger - c^0|p')$ commutes with the total Lagrangian. Therefore we can reduce the interaction Lagrangian L_I to

$$L_I = -c\,(p|j^k|p')\,(p'|b_k + b_k^\dagger|p)$$
$$= -c\,(p|j^k|p')\,(p'|b_k|p) - c\,(p|j^k|p')\,(p'|b_k^\dagger|p) \tag{26}$$

imposing the Lorentz condition on state vectors

$$(p|[p^k, A_k]|p')\,\Psi = 0,$$
$$\Phi\,(p|[p^k, A_k]|p') = 0,$$

or, preserving only the subspaces labeled by the eigenvalues 0 of the operators $(p|[p^k, A_k]|p')$.

The main task of quantum electrodynamics is to bring the total Lagrangian into

diagonal form. There is an operator that commutes with the total Lagrangian L. The operator is the total number of electrons $N = (\phi_\alpha|p)\,(p|\psi'_\alpha)$. As is easily verified,

$$[N, (p|j^k|p')] = \gamma^k_{\alpha\beta}[(\phi_\tau|p'')(p''|\psi_\tau),\ (\phi_\alpha|p')(p|\psi_\beta)]$$

$$= \gamma^k_{\alpha\beta}\{(\phi_\tau|p'')\,\delta_{\tau\alpha}(p''|p')\,(p|\psi_\beta) - (\phi_\alpha|p')\,\delta_{\beta\tau}(p|p'')\,(p''|\psi_\tau)\}\ \text{from (14)}$$

$$= (\phi_\alpha|p')\gamma^k_{\alpha\beta}(p|\psi_\beta) - (\phi_\alpha|p')\gamma^k_{\alpha\beta}(p|\psi_\beta) = 0.$$

N commutes with each component of the current since the current contains a pair of creation and annihilation operators. Hence $[L, N] = 0$.

Therefore the eigenspaces of the total Lagrangian can be classified into independent subspaces by the eigenvalues of N. This is an important property.

When the eigenvalue of N is equal to 0, that is, no electron quantum is present in the entire space-time, the total Lagrangian is reduced to L_R alone, since each component of the current becomes a null operator. There exist then only photon quanta, real as well as virtual, in the entire space-time. They are not cursed with their self-energy.

§ 12. The application of the perturbation method

It can hardly be justified to use the perturbation method in diagonalizing the total Lagrangian L, which has infinite degrees of freedom. Nevertheless we are forced to use the method in order to compare the consequences of our formulation with those of the Hamiltonian formulation in which the method is used to deduce its consequences.

The perturbation method applied to quantum electrodynamics gets clearer with the use of Feynman diagrams.[11]

Since the number of electrons must be invariant throughout every process of interaction with photons, the polygon representing an electron, irrespective of its energy, must extend infinitely in both ends, so that the loop of electron line never appears.

When no electron is present, that is, $N = 0$, a permissible diagram consists of any number of photon-lines. The electron-loop is not permissible.

When an electron is present, that is, $N = 1$, permissible diagrams up to the second order in e are those of pair creation, pair annihilation, Compton scattering and self-energy.

We shall compute the matrix elements for the above processes.

Firstly we eliminate the term depending on e, that is, the interaction Lagrangian L_I from the total Lagrangian L by a canonical transformation

$$e^{-X}Le^{X} = L + [L, X] + 1/2 \cdot [[L, X], X] + \cdots$$

$$= L_0 + (L_I + [L_0, X]) + ([L_I, X] + 1/2 \cdot [[L_0, X]X]) + \cdots$$

with
$$L_0 = L_M + L_R.$$

We determine X such that

$$L_I + [L_0, X] = 0. \tag{27}$$

Then we have

$$e^{-X}Le^{X} = L_0 + 1/2 \cdot [L_I, X] + \cdots$$

up to the second order in e.

Since L_M and L_R are bilinear in v, u and $b^\dagger$, b respectively, we can put

$$X = e\,(p'|\xi^k|p)\,(p|b_k|p') + e\,(p'|\eta^k|p)\,(p|b_k{}^\dagger|p'),$$

where

$$(p'|\xi^k|p) = (v_\alpha|p)\,\xi^k_{\alpha\beta}(p,p')\,(p'|u_\beta)$$

$$(p'|\eta^k|p) = (v_\alpha|p)\,\eta^k_{\alpha\beta}(p,p')\,(p'|u_\beta)$$

are quantities to be determined by the equation (27) with the interaction (26).

$\xi^k_{\alpha\beta}$, $\eta^k_{\alpha\beta}$ are found to be

$$\left.\begin{aligned}
\xi^k_{\alpha\beta}(p,p') &= \alpha^k_{\alpha\beta}(p,p') / \{q_\alpha(p) - q_\beta(p') - 2(p-p')^2\} \\
\eta^k_{\alpha\beta}(p,p') &= \alpha^k_{\alpha\beta}(p,p') / \{q_\alpha(p) - q_\beta(p') + 2(p-p')^2\}.
\end{aligned}\right\} \quad (28)$$

(see Appendix 1)

Secondly we compute the commutator $[L_I, X]$.

$$\left.\begin{aligned}
[L_I, X] = &-e^2[(p'|j^k|p),\ (p'''|\xi^j|p'')]\,(p|b_k|p')\,(p''|b_j|p''') & (A) \\
&-e^2[(p'|j^k|p),\ (p'''|\eta^j|p'')]\,(p|b_k{}^\dagger|p')\,(p''|b_j{}^\dagger|p''') & (B) \\
&-e^2[(p'|j^k|p),\ (p'''|\eta^j|p'')]\,(p''|b_k{}^\dagger|p''')\,(p|b_k|p') \\
&-e^2[(p'|j^k|p),\ (p'''|\xi^j|p'')]\,(p|b_k{}^\dagger|p')\,(p''|b_j|p''') & (C) \\
&-e^2(p'|j^k|p)\,(p'''|\eta^j|p'')[(p|b_k|p'),(p''|b_j{}^\dagger|p''')] \\
&-e^2(p'''|\xi^j|p'')\,(p'|j^k|p)[(p|b_k{}^\dagger|p'),(p''|b_j|p''')]. & (D)
\end{aligned}\right\} \quad (29)$$

(see Appendix 2)

Using commutation relations (17), (23), we have

$$(A) = e^2\{\xi^k_{\alpha\tau}(p,p'')\,\alpha^j_{\tau\beta}(p'',p') - \alpha^k_{\alpha\tau}(p,p'')\,\xi^j_{\tau\beta}(p'',p')\} \cdot$$
$$(v_\alpha|p)\,(p'|u_\beta)\,(p|b_k|p'')\,(p''|b_j|p'),$$

$$(B) = e^2\{\eta^k_{\alpha\tau}(p,p'')\,\alpha^j_{\tau\beta}(p'',p') - \alpha^k_{\alpha\tau}(p,p'')\,\eta^j_{\tau\beta}(p'',p')\} \cdot$$
$$(v_\alpha|p)\,(p'|u_\beta)\,(p|b_k{}^\dagger|p'')\,(p''|b_j{}^\dagger|p'),$$

$$(C) = e^2\{\eta^k_{\alpha\tau}(p,p'')\,\alpha^j_{\tau\beta}(p'',p') - \alpha^k_{\alpha\tau}(p,p'')\,\xi^j_{\tau\beta}(p'',p')\} \cdot$$
$$(v_\alpha|p)\,(p'|u_\beta)\,(p|b_k{}^\dagger|p'')\,(p''|b_j|p')$$

$$+ e^2\{\xi^k_{\alpha\tau}(p,p'')\,\alpha^j_{\tau\beta}(p'',p') - \alpha^k_{\alpha\tau}(p,p'')\,\eta^j_{\tau\beta}(p'',p')\} \cdot$$
$$(v_\alpha|p)\,(p'|u_\beta)\,(p''|b_j{}^\dagger|p')\,(p|b_k|p''),$$

$$(D) = e^2\{\eta^j_{\tau\delta}(p',p) - \xi^j_{\tau\delta}(p',p)\}\,\alpha^k_{\alpha\beta}(p,p')\,(v_\tau|p')\,(v_\alpha|p)\,(p|u_\delta)\,(p'|u_\beta) \cdot$$
$$\left(g_{jk} - \frac{(p_j - p_j')(p_k - p_k')}{(p-p')^2}\right) \quad (D')$$

$$+ e^2\{\xi^j_{\alpha\tau}(p,p')\,\alpha^k_{\tau\beta}(p',p) - \alpha^k_{\alpha\tau}(p,p')\,\eta^j_{\tau\beta}(p',p)\}\,(p'|p')\,(v_\alpha|p)\,(p|u_\beta) \cdot$$

$$\left(g_{jk}-\frac{(p_j-p_j')\,(p_k-p_k')}{(p-p')^2}\right)\,. \quad (D'')$$

Finally we eliminate field variables referring to virtual quanta, leaving only field variables referring to real quanta so that we may get the effective Lagrangian for the transition from an initial state to a final state through an intermediate state. The Lagrangian for free fields L_0 is then reduced to nothing because the respective Lagrangians $q_\alpha(p)$, $(p-p')^2$ of electron and photon vanish for real quanta.

Hence we have

$$e^{-X}L e^{X}=\frac{1}{2}[L_I,\,X]=\frac{1}{2}\,(A)+\frac{1}{2}\,(B)+\frac{1}{2}\,(C)+\frac{1}{2}\,(D')+\frac{1}{2}(D'')$$

$$\frac{1}{2}\,(A)=e^2(\phi|p)\gamma^k\frac{1}{-p''+m}\gamma^j(p'|\psi)\,(p|b_k|p'')\,(p''|b_j|p') \qquad (\alpha)$$

$$\frac{1}{2}\,(B)=e^2(\phi|p)\gamma^k\frac{1}{-p''+m}\gamma^j(p'|\psi)\,(p|b_k^\dagger|p'')\,(p''|b_k^\dagger|p') \qquad (\beta)$$

$$\frac{1}{2}\,(C)=e^2(\phi|p)\gamma^k\frac{1}{p''-m}\gamma^j(p'|\psi)\,\{(p|b_k^\dagger|p'')\,(p''|b_j|p')+(p''|b_j^\dagger|p')\,(p|b_k|p'')\} \qquad (\gamma)$$

$$\frac{1}{2}(D')=e^2\gamma_{\tau\delta}^j\gamma_{\alpha\beta}^k\frac{1}{2\,(p-p')^2}(\phi_\tau|p')\,(\phi_\alpha|p)(p|\psi_\delta)\,(p'|\psi_\beta)\left(g_{jk}-\frac{(p_j-p_j')(p_k-p_k')}{(p-p')^2}\right) \qquad (\delta)$$

$$\frac{1}{2}(D'')=-e^2(\phi|p)\gamma^k\frac{(p'|p')}{p'-m+2(p-p')^2}\gamma^j(p|\psi)\left(g_{jk}-\frac{(p_j-p_j')\,(p_k-p_k')}{(p-p')^2}\right)(\varepsilon)\,.$$

(30)

(see Appendix 3)

(α), (β), (γ) give exactly the matrix elements for pair creation, pair annihilation, Compton scattering respectively. (δ) gives the matrix element for Møller scattering except for a factor.

(ε) gives the electron self-energy, which is different from the usual one. The denominator $p'-m+2(p-p')^2$ is the sum of two Lagrangians $p'-m$, $2(p-p')^2$ instead of their product in the usual formulation. The term $|p'-m+2(p-p')^2|^{-1}$ $(p'|p')$, which is equal to

$$\int\frac{d^4p'}{p'-m+2(p-p')^2}\,,$$

has no definite value, so we fail to eliminate the electron self-energy. The term referring to photon self-energy, however, never appears in agreement with the preceding remark at the end of last section.

The writer expresses hearty thanks to MM. K. Aizu and T. Okabayashi for their kind advices and criticisms.

Appendix 1, Note to (28)

To obtain these relations, we compute at first

$$[L_M, (p'|\xi^k|p)] = q_\tau(p'')\xi^k_{\alpha\beta}(p,p')[(v_\tau|p'')(p''|u_\tau),(v_\alpha|p)(p'|u_\beta)] \quad \text{from (16)}$$

$$= q_\tau(p'')\xi^k_{\alpha\beta}(p,p')[(v_\tau|p'')\{(p''|u_\tau),(v_\alpha|p)\}(p'|u_\beta)$$
$$- (v_\alpha|p)\{(p'|u_\beta),(v_\tau|p'')\}(p''|u_\tau)]$$

$$= q_\tau(p'')\xi^k_{\alpha\beta}(p,p')\{(v_\tau|p'')\partial_{\tau\alpha}(p''|p)(p'|u_\beta)$$
$$- (v_\alpha|p)\partial_{\beta\tau}(p'|p'')(p''|u_\tau)\} \quad \text{from (17)}$$

$$= q_\alpha(p)\xi^k_{\alpha\beta}(p,p')(v_\alpha|p)(p'|u_\beta) - q_\beta(p')\xi^k_{\alpha\beta}(p,p')(v_\alpha|p)(p'|u_\beta)$$

$$= \{q_\alpha(p) - q_\beta(p')\}(v_\alpha|p)\xi^k_{\alpha\beta}(p,p')(p'|u_\beta)$$

and secondly

$$[L_R, (p|b_k|p')] = 2(p''-p''')^2[(p''|b^\dagger_i|p''')(p'''|b^i_k|p''),(p|b_k|p')] \quad \text{from (24)}$$

$$= -2(p''-p''')^2[(p|b_k|p'),(p''|b^\dagger_i|p''')](p'''|b^i_k|p'')$$

$$= -2(p''-p''')^2\Big(g_{ki} - \frac{(p_k-p'_k)(p_i-p'_i)}{(p-p')^2}\Big)(p|p''')(p''|p')(p'''|b^i_k|p'')$$
$$\text{from (23)}$$

$$= -2(p-p')^2\Big(g_{ki} - \frac{(p_k-p'_k)(p_i-p'_i)}{(p-p')^2}\Big)(p|b^i_k|p')$$

$$= -2(p-p')^2(p|b_k|p') \quad \text{from (22)}$$

so we get

$$[L_M + L_R, (p'|\xi^k|p)(p|b_k|p')]$$
$$= \{q_\alpha(p) - q_\beta(p') - 2(p-p')^2\}(v_\alpha|p)\xi^k_{\alpha\beta}(p,p')(p'|u_\beta)(p|b_k|p')$$

which is to cancel the first half of the interaction Lagrangian (26)

$$-(p'|j^k|p)(p|b_k|p') = -(v_\alpha|p)a^k_{\alpha\beta}(p,p')(p'|u_\beta)(p|b_k|p').$$

So we get the first relation of (28). To obtain the second, it is sufficient to note that

$$[L_R, (p|b^\dagger_k|p')] = 2(p-p')^2(p|b^\dagger_k|p').$$

Appendix 2, Note to (29)

The following equality is used

$$[A_1B_1 + A_2B_2, A_3B_3 + A_4B_4] = [A_1, A_3]B_1B_3 + [A_2, A_4]B_2B_4 + [A_1, A_4]B_4B_1$$
$$+ [A_2, A_3]B_2B_3 + A_1A_4[B_1, B_4] + A_3A_2[B_2, B_3],$$

where

$$[A_i, B_j] = 0, \quad [B_1, B_3] = [B_2, B_4] = 0.$$

Appendix 3, Note to (30)

The assumption that remaining field variables refer to real quanta alone gives

$$\xi^k_{\alpha\tau}(p,p'') = \frac{\alpha^k_{\alpha\tau}(p,p'')}{q_\alpha(p) - q_\tau(p'') - 2(p-p'')^2} = \frac{\alpha^k_{\alpha\tau}(p,p'')}{-q_\tau(p'')}$$

$$\xi^j_{\tau\beta}(p'',p') = \frac{\alpha^j_{\tau\beta}(p'',p')}{q_\tau(p'') - q_\beta(p') - 2(p''-p')^2} = \frac{\alpha^j_{\tau\beta}(p'',p')}{q_\tau(p'')}$$

$$\eta^k_{\alpha\tau}(p,p'') = \frac{\alpha^k_{\alpha\tau}(p,p'')}{q_\alpha(p) - q_\tau(p'') + 2(p-p'')^2} = \frac{\alpha^k_{\alpha\tau}(p,p'')}{-q_\tau(p'')}$$

$$\eta^j_{\tau\beta}(p'',p') = \frac{\alpha^j_{\tau\beta}(p'',p')}{q_\tau(p'') - q_\beta(p') + 2(p''-p')^2} = \frac{\alpha^j_{\tau\beta}(p'',p')}{q_\tau(p'')}$$

because of the vanishing of Lagrangians

1)　$(p-p'')^2$, $(p''-p')^2$, which refer to real photons $(p|b_k|p'')$, $(p''|b_j|p')$

2)　$q_\alpha(p)$, $q_\beta(p')$, which refer to real electrons $(v_\alpha|p)$, $(p'|u_\beta)$.
Then (A) reduces to

$$-2c^2 \alpha^k_{\alpha\tau}(p,p'') \frac{1}{q_\tau(p'')} \alpha^j_{\tau\beta}(p'',p')(v_\alpha|p)(p'|u_\beta)(p|b_k|p'')(p''|b_j|p').$$

Further we have

$$(v_\alpha|p)\alpha^k_{\alpha\tau}(p,p'') \frac{1}{q_\tau(p'')} \alpha^j_{\tau\beta}(p'',p')(p'|u_\beta)$$

$$= (v_\alpha|p)K^{-1}_{\alpha\delta}(p)\gamma^k_{\delta\epsilon}K_{\epsilon\tau}(p'') \frac{1}{q_\tau(p'')} K^{-1}_{\tau\rho}(p'')\gamma^j_{\rho\sigma}K_{\sigma\beta}(p')(p'|u_\beta) \qquad \text{from (25)}$$

$$= (\phi_\delta|p)\gamma^k_{\delta\epsilon}\left(\frac{1}{p''-m}\right)_{\epsilon\rho} \gamma^j_{\rho\sigma}(p'|\psi_\sigma) \qquad \text{from (15)}$$

$$= (\phi|p)\gamma^k \frac{1}{p''-m}\gamma^j(p'|\psi),$$

so we get

$$(A) = -2c^2(\phi|p)\gamma^k \frac{1}{p''-m}\gamma^j(p'|\psi)(p|b_k|p'')(p''|b_j|p'), \text{ etc.}$$

References

1)　W. Heisenberg and W. Pauli, ZS. f. Phys. **56**(1929), 1.
2)　S. Tomonaga, Prog. Theor. Phys. **1**(1946), 27.
3)　J. Schwinger, Phys. Rev. **74**(1948), 1439.
4)　G. Iwata, Prog. Theor. Phys. **6** (1951), 684.
5)　　　　"　　　　　"　　　　　**7** (1952), 39.
6)　C. Gregory, Phys. Rev. **78** (1950), 479.
7)　H. Yukawa, Prog. Theor. Phys. **2**(1947), 209.
8)　P. Kristensen and C. Møller, Det. Dansk. Vid. Selsk. **27**(1952) No. 7.
9)　H. Weyl, *Classical Groups* (Princeton 1946), 173.
10)　P. Ehrenfest, ZS. f. Phys. **45**(1927), 455.
11)　R. P. Feynman, Phys. Rev. **76**(1949), 769.

The Theory of Collisions between Two Diatomic Molecules

Kazuo TAKAYANAGI

Department of Physics, Saitama University

(Received March 8, 1954)

The general features of the elastic and the inelastic collisions have been investigated in this paper. The inelastic collisions are confined to the rotational and the vibrational ones, and the electronic degrees of freedom are disregarded throughout. Starting from the fundamental set of equations for the collision problem, the effects of the non-spherical character of the intermolecular potential are discussed. For the rotational transitions where the level spacings are rather small, the semi-classical method is employed and for the vibrational transitions, the distorted wave method is developed and some simple methods to estimate the cross sections are also discussed. Special considerations are also given for the cases in which the two colliding molecules are identical.

§ 1. Introduction

As is well known, the knowledge of molecular collision processes, in general, together with that of the stationary states of molecules form the fundamental part of the chemical physics. And these molecular collisions which may be accompanied with the rotational and/or the vibrational transitions can be treated, in principle, by solving the well-established Schroedinger wave equation for that system. The straightforward solution, however, impels us to carry out the extremely laborious calculations. It may be said that it is almost impossible to do so unless some specially designed calculating machines are invented. For this reason, the calculations for the elastic molecular collisions are usually done by assuming the spherically symmetric potential without taking the internal degrees of freedom into account, while the calculations of the probabilities of the vibrational or the rotational transitions have been made for the special orientations of the two colliding molecules (e.g. collinear collision or collision in a plane).[1]–[3] Thus, it is desirable to investigate more fully the general feature of the collision processes and to seek for some approximation methods of treatment. In this paper, some detailed investigations for the collisions between two diatomic molecules in the unrestricted space are reported.

First thing we have to do is to set up the fundamental equations for our problem. Then we must determine the intermolecular force for particular case and calculate its matrix elements which appear in the fundamental equations. The pure theoretical determination of the intermolecular potential presents one of the outstanding difficulties for the collision problem, and the semi-empirical determination is preferable if it is possible. These problems are investigated in the first part (§ 2, § 3). Special considerations are also given for the case in which the two colliding molecules are identical. (§ 4) In the next part

of this paper, the cross section for the elastic scattering is defined in relation to the asymptotic behaviour of the wave function ($\S 5$), and some considerations, especially on the effects of the non-spherical nature of the potential, are given for the elastic scattering. ($\S 6$)   The inelastic transitions, especially the vibrational one, in the thermal collisions may be treated by the distorted wave method. ($\S 7$)   The general formulation on this line and some attempts to estimate the cross sections in simple ways are explained in $\S 9$.   For the rotational transitions, however, the transition probabilities are rather large so that the distorted wave method can no longer be employed.   Instead, the semi-classical method, in which the translational motion is described classically, is proposed for this case. ($\S 8$) Finally, we must consider the statistical problems, such as the viscosity, the diffusion, the ultrasonic dispersion and so forth, in order to combine the theory of elementary processes with the experimental data.   In particular, we are interested in the effects due to the existence of the internal degrees of freedom.   Some brief discussions are given for these effects. ($\S 10$)

The numerical examples are not shown in this paper.   It is also desirable to investigate the other types of collision, such as the collisions between the larger molecules and the multiple collisions in which more than two molecules collide simultaneously.   These problems are postponed to another chance.

§ 2.   Derivation of the fundamental set of equations[*]

We shall first consider the wave function of a free diatomic molecule.   As is well-known, it has the following form

$$u_{nl}(\bar{\xi}) Y_{lm}(\theta, \varphi)/\bar{\xi},$$

where $\bar{\xi}$ is the internuclear distance, and θ, φ define the direction of the molecular axis with respect to the fixed coordinate axes in the space. u_{nl} is the vibrational wave function which satisfies the equation of the form[**]

$$\{d^2/d\bar{\xi}^2 - l(l+1)/\bar{\xi}^2 - 2\mu V(\bar{\xi}) + 2\mu W_{nl}\} u_{ln}(\bar{\xi}) = 0,$$

together with the boundary conditions ($u(0)=0$ and $u(\bar{\xi}) \to 0$ for $\bar{\xi} \to \infty$), and is normalized according to

$$\int_0^\infty u_{nl}^*(\bar{\xi}) u_{n'l}(\bar{\xi}) d\bar{\xi} = \delta_{nn'}.$$

Y_{lm} is the ordinary spherical harmonic function :

$$Y_{lm}(\theta, \varphi) = \Theta_{lm}(\theta) \Phi_m(\varphi),$$

$$\left.\begin{array}{l} \Theta_{lm}(\theta) = (-1)^m \left\{ \dfrac{2l+1}{2} \dfrac{(l-m)!}{(l+m)!} \right\}^{1/2} \sin^m \theta \dfrac{d^m}{(d\cos\theta)^m} P_l(\cos\theta), \\[3mm] \Theta_{l-m}(\theta) = (-1)^m \cdot \Theta_{lm}(\theta), \end{array}\right\} \quad (m > 0)$$

---

[*] Essential part of $\S\S$ 1–3 was reported in the international symposium on molecular physics at Nikko, Japan (September, 1953).

[**] Atomic units are used throughout this paper.

$$\Phi_m(\varphi) = (1/\sqrt{2\pi})\exp(im\varphi).$$

$V(\xi)$ in the above equation is the vibrational potential, and μ means the reduced mass of the two atoms. W_{nl} is the internal energy of the molecule, disregarding the electronic one. The electronic wave function does not appear explicitly in this work.

Now we shall proceed to the system consisting of the two diatomic molecules. If the two molecules are far apart, the potential energy of the system is simply the sum of the intramolecular potentials of both molecules

$$V = V_1(\xi_1) + V_2(\xi_2),$$

and when these two molecules approach to each other, the mutual interaction starts to act and the potential energy function may be written as

$$V = V_1(\xi_1) + V_2(\xi_2) + V'(R, \theta, \Phi, \xi_1, \theta_1, \varphi_1, \xi_2, \theta_2, \varphi_2). \qquad (2\cdot1)$$

(See Fig. 1) Since the third term, intermolecular potential, depends on the internal coordinates of the two molecules, the rotational transition (change of l, m) and the vibrational transition (change of n) can take place. The separation of the potential into three parts as in the expression $(2\cdot1)$ is useful only in the distant collision such as the thermal collision in the room
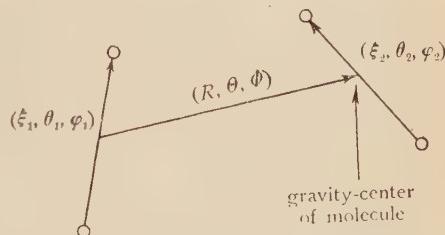
Fig. 1

temperature, and may become meaningless in the extremely close collision, especially in the case of the rearrangement chemical reaction such as the double decomposition : $AB + CD = AC + BD$, which is beyond the scope of our present considerations.

On the other hand, the kinetic energy of this system may be expressed as

$$T = -(1/2\mu_1)\nabla_1^2 - (1/2\mu_2)\nabla_2^2 - (1/2M)\nabla_R^2, \qquad (2\cdot2)$$

where M means the reduced mass of the two-molecule-system. The first and the second terms in the right hand side correspond to the internal motions of two molecules and the third term corresponds to the relative motion of them. The translational energy of the system as a whole has been separated out, which is always possible without loss of generality.

Then the Schroedinger wave equation is given by

$$(T + V - E)\Psi = 0 \qquad (2\cdot3)$$

where Ψ is the wave function representing the whole system.

Now if the intermolecular potential is small, or if, at least, the dependence of the intermolecular potential on the internal coordinates of two molecules is assumed to be little, it is convenient to expand the total wave function as follows :

$$\Psi = \sum_\alpha (1/R) f_\alpha(R) A_\alpha(\theta, \Phi, \xi_1, \theta_1, \varphi_1, \xi_2, \theta_2, \varphi_2). \qquad (2\cdot4)$$

$$A_\alpha = Y_{j\nu}(\theta\Phi)\frac{1}{\xi_1}u_{nl}(\xi_1)Y_{lm}(\theta_1\varphi_1)\frac{1}{\xi_2}u_{n'l'}(\xi_2)Y_{l'm'}(\theta_2\varphi_2),$$

where α stands for $(j, \nu, n, l, m, n', l', m')$. By substituting this expansion into $(2\cdot3)$, we have a set of equations for $f(R)$:

$$\frac{1}{2M}\left\{\frac{d^2}{dR^2}-\frac{j(j+1)}{R^2}+k_\alpha{}^2\right\}f_\alpha(R)=\sum_{\alpha'}(\alpha|V'|\alpha')f_{\alpha'}(R), \qquad (2\cdot5)$$

where

$$k_\alpha{}^2/2M=E-W'^{(1)}_{nl}-W'^{(2)}_{n'l'}, \qquad (2\cdot6)$$

$$(\alpha|V'|\alpha')=\iiint\int d\Omega\, d\Omega_1 d\Omega_2 \xi_1{}^2 d\xi_1 \xi_2{}^2 d\xi_2 A_\alpha\cdot V'\cdot A_{\alpha'}, \qquad (2\cdot7)$$

$$d\Omega=\sin\theta\, d\theta\, d\Phi, \qquad d\Omega_i=\sin\theta_i d\theta_i d\varphi_i.$$

In the above treatment, the total angular momentum of the system is not definite. However, since the total angular momentum quantum numbers J, μ (μ corresponds to the z-component of the angular momentum) are the good quantum numbers, it is desirable to use the representation in which the values of these quantum numbers appear explicitly. For this sake, the following linear combinations of the single products of the spherical harmonics must be employed as the basic functions of the expansion:

$$\boldsymbol{Y}_{jLll'}^{J\mu}=\sum_{m,m'}s^{jL}_{J\mu-m\,m}s^{Ll'}_{L\,m'\,m-m'}Y_{j\,\mu-m}(\theta\Phi)Y_{lm}(\theta_1\varphi_1)Y_{l'\,m-m'}(\theta_2\varphi_2). \qquad (2\cdot8)^{*)}$$

Then the equations $(2\cdot4)$, $(2\cdot5)$ and $(2\cdot7)$ are respectively replaced by

$$\Psi=\sum_\beta(1/R)F_\beta(R)\cdot B_\beta(\theta,\,\Phi,\,\xi_1,\,\theta_1,\,\varphi_1,\,\xi_2,\,\theta_2,\,\varphi_2) \qquad (2\cdot9)$$

$$B_\beta\equiv(1/\xi_1)u_{nl}(\xi_1)(1/\xi_2)u_{n'l'}(\xi_2)\,\boldsymbol{Y}_{jLll'}^{J\mu},$$

$$\frac{1}{2M}\left\{\frac{d^2}{dR^2}-\frac{j(j+1)}{R^2}+k_\beta{}^2\right\}F_\beta(R)=\sum_{\beta'}(\beta|V'|\beta')F_{\beta'}(R), \qquad (2\cdot10)$$

$$\beta\equiv(J,\,\mu,\,n,\,n',\,j,\,L,\,l,l'),$$

$$(\beta|V'|\beta')=\delta_{J,J'}\delta_{\mu,\mu'}\iiint\int d\Omega\, d\Omega_1 d\Omega_2 \xi_1{}^2 d\xi_1 \xi_2{}^2 d\xi_2 B_\beta\cdot V'\cdot B_{\beta'}. \qquad (2\cdot11)$$

Sometimes, the nuclear and/or electronic spin may exist which have not been taken into account in the above equations. But the spin does not couple with the orbital angular momenta strongly, J and μ are still good quantum numbers even if we don't include the spin in this resultant angular momentum.

The equations $(2\cdot10)$ are our fundamental set of equations to be solved. The next step is the determination of the matrix elements in these equations ($\S$ 3), and then comes the investigation of the general features of the solution ($\S\S$ 6–9).

If we can neglect the vibrational motion of the molecules, $u(\xi)$ is replaced by

$$\sqrt{\delta(\xi-\xi_0)}$$

where ξ_0 means the equilibrium value of the internuclear distance.

---

* The so-called Clebsch-Gordan's coefficients $s^{ll'}_{lmm'}$ is the same with that defined by Wigner in his text book.[7] See also Appendix II.

Recently, Hirschfelder, Curtiss and Adler in Wisconsin University have also investigated the general formulation of the collision problem for the polyatomic molecules.[4)–6)] They have separated the rotational coordinates from the N-particle Schroedinger equation by making use of the group theoretical method and have obtained[6)] a similar set of equations to our equations $(2 \cdot 10)$.

§ 3. Intermolecular force potential and its matrix elements

All kinds of collision process have its origin in the intermolecular potential V'', and usually the cross sections of various processes depend rather sensitively on this potential function. But, unfortunately, the calculation of the intermolecular force potential is a very difficult thing, and no theoretical calculations of V'' have been made for diatomic molecules, except for H_2–H_2 force.[8)–10)] For this reason, it is preferable to use the semi-empirical formula as far as possible. For instance, we may determine the potential function from the equation of state of the gas or from some transport phenomena such as viscosity and then employ the so-determined function to calculate the cross sections of some inelastic processes.

The intermolecular force is usually divided into two parts. The one is the strong repulsive force (V_r) in the closest distances which arises from the overlapping of the electron clouds of the two molecules. The other is the so-called Van der Waals force (V_a) which is usually attractive and have a long range of action. The former may be expressed approximately by the sum of the interatomic repulsive potentials between the atoms belonging to the different molecules, i.e.

$$V_r = v_{13}(r_{13}) + v_{14}(r_{14}) + v_{23}(r_{23}) + v_{24}(r_{24}), \qquad \qquad (3 \cdot 1)$$

where r_{ij} is the distance between i-th and j-th atoms.[9) 11)] (See Fig. 2.) The interatomic potential v_{ij} may be determined empirically, by assuming a simple form containing some parameters which should be adjusted by comparing with some experimental data.

If the internuclear distances ξ_1, ξ_2 are both small compared with the intermolecular distance R (this is roughly true for the thermal collisions at ordinary temperatures), it is convenient to expand the function $(3 \cdot 1)$ into the following series

Fig. 2

$$V_r = \sum_{a=1,2} \sum_{b=3,4} \sum_{i,j,k} \sum_{l,m,n}^{\infty} \sum^{\infty} \frac{x_a^i y_a^j z_a^k}{i!\,j!\,k!} \frac{x_b'^l y_b'^m z_b'^n}{l!\,m!\,n!} \left[\frac{\partial^{i+j+k+l+m+n} v_{ab}(r)}{\partial X^i \partial Y^j \partial Z^k \partial X'^l \partial Y'^m \partial Z'^n} \right]_{\substack{X=Y=Z=0 \\ X'=Y'=Z'=0}}$$

$$(r^2 = (X'-X-R)^2 + (Y'-Y)^2 + (Z'-Z)^2)$$

$$= \{ v_{13}(R) + v_{14}(R) + v_{23}(R) + v_{24}(R) \}$$
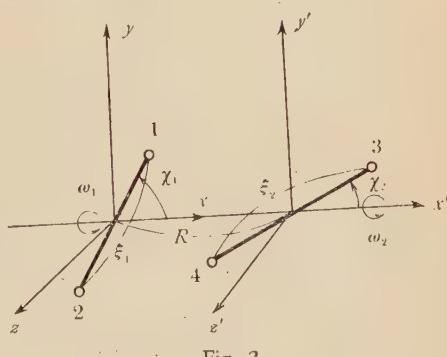
$$+ \{(x_1-x_3')(\partial v_{13}/\partial X)_0 + (x_1-x_4')(\partial v_{14}/\partial X)_0$$

$$+ (x_2-x_3')(\partial v_{23}/\partial X)_0 + (x_2-x_4')(\partial v_{24}/\partial X)_0\}$$

$$+ \left[\tfrac{1}{2}\{(x_1^2+x_3'^2)(\partial^2 v_{13}/\partial X^2)_0 + (x_1^2+x_4'^2)(\partial^2 v_{14}/\partial X^2)_0\right.$$

$$+ (x_2^2+x_3'^2)(\partial^2 v_{23}/\partial X^2)_0 + (x_2^2+x_4'^2)(\partial^2 v_{24}/\partial X^2)_0$$

$$+ (y_1^2+y_3'^2+z_1^2+z_3'^2)(\partial^2 v_{13}/\partial Y^2)_0 + (y_1^2+y_4'^2+z_1^2+z_4'^2)(\partial^2 v_{14}/\partial Y^2)_0$$

$$+ (y_2^2+y_3'^2+z_2^2+z_3'^2)(\partial^2 v_{23}/\partial Y^2)_0 + (y_2^2+y_4'^2+z_2^2+z_4'^2)(\partial^2 v_{24}/\partial Y^2)_0\}$$

$$+ \{x_1 x_3'(\partial^2 v_{13}/\partial X \partial X')_0 + x_1 x_4'(\partial^2 v_{14}/\partial X \partial X')_0$$

$$+ x_2 x_3'(\partial^2 v_{23}/\partial X \partial X')_0 + x_2 x_4'(\partial^2 v_{24}/\partial X \partial X')_0$$

$$+ (y_1 y_3'+z_1 z_3')(\partial^2 v_{13}/\partial Y \partial Y')_0 + (y_1 y_4'+z_1 z_4')(\partial^2 v_{14}/\partial Y \partial Y')_0$$

$$+ (y_2 y_3'+z_2 z_3')(\partial^2 v_{23}/\partial Y \partial Y')_0 + (y_2 y_4'+z_2 z_4')(\partial^2 v_{24}/\partial Y \partial Y')_0\}\Big]$$

$+ 3rd$ *and higher order terms.*

$$(3 \cdot 2)$$

It is easy to see that this expression is a function of three cosines $\cos \chi_1$, $\cos \chi_2$ and $\cos \chi'$ (See Fig. 3) and ξ_1, ξ_2, R. This can be proved either by substituting the relations

$$x_1 = \xi_1(m_2/m_1+m_2)\cos\chi_1,$$

$$(m_i \text{ is the mass of the } i\text{-th atom.})$$

$$y_1^2+z_1^2 = \xi_1^2(m_2/m_1+m_2)^2(1-\cos^2\chi_1),$$

$$y_1 y_3'+z_1 z_3' = \xi_1 \xi_2(m_2/m_1+m_2)$$

$$\times (m_4/m_3+m_4)\sin\chi_1 \sin\chi_2 \cos(\omega_1-\omega_2)$$

$$= \xi_1 \xi_2(m_2/m_1+m_2)(m_4/m_3+m_4)$$

$$\times (\cos\chi'-\cos\chi_1 \cos\chi_2) \text{ etc.,}$$
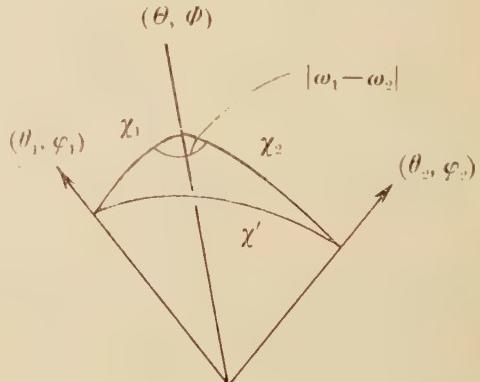
Fig. 3

into the above expression directly, or by noticing that these three cosines together with ξ_1, ξ_2 and R determine the distances of each atom from three other atoms uniquely. Thus, after all, the repulsive potential V_r must have the form

$$V_r = \sum_{\alpha\beta\gamma} v_{\alpha\beta\gamma}^{(r)}(\xi_1, \xi_2, R) P_\alpha(\cos\chi_1) P_\beta(\cos\chi_2) P_\gamma(\cos\chi') \tag{3·3}$$

and, if the anisotropy of this potential is not so large, only the first few terms have to be employed.

Now it is easy to calculate the matrix elements of V_r. By remembering the addition theorem

$$P_\alpha(\cos\chi_1) = (4\pi/2\alpha+1)\sum_{\mu=-\alpha}^{\alpha} Y_{\alpha\rho}^*(\theta\Phi) Y_{\alpha\rho}(\theta_1\varphi_1)$$

$$P_\beta(\cos \chi_2) = (4\pi/2\beta+1) \sum_{\sigma=-\beta}^{\beta} Y_{\beta\sigma}{}^*(\Theta\Phi) Y_{\beta\sigma}(\theta_2\varphi_2)$$

$$P_\gamma(\cos \chi') = (4\pi/2\gamma+1) \sum_{\tau=-\gamma}^{\gamma} Y_{\gamma\tau}{}^*(\theta_1\varphi_1) Y_{\gamma\tau}(\theta_2\varphi_2), \tag{3.4}$$

it is clear that the matrix elements of the function (3.3) can be expressed in terms of

$$c^L(l'\, m'\,;\, lm) = (2/2L+1)^{1/2} \iint Y_{l'm'}{}^*(\theta\varphi) Y_{Lm'-m}(\theta\varphi) Y_{lm}(\theta\varphi) \sin\theta\, d\theta\, d\varphi.$$

The values of these integrals have been calculated for not too large values of l and l'.[12) 13)] For the larger values of l and l', we have to calculate by making use of the general formula for these integrals.[12) 14)]

The matrix elements of the Van der Waals force V_a can be derived similarly if this potential has been expressed in a similar form to (3.3).

If two molecules are homonuclear, we have $x_1=-x_2,\ y_1=-y_2$ etc., and furthermore the four interatomic potentials $v_{13}, v_{14}, v_{23}, v_{24}$ become same functions and, therefore, the repulsive potential can be expressed in a much simpler form:

$$V_r=4v(R)+2\left[(x_1^2+x_3'^2)(\partial^2 v/\partial X^2)_0+(y_1^2+y_3'^2+z_1^2+z_3'^2)(\partial^2 v/\partial Y^2)_0\right], \tag{3.5}$$

higher terms being neglected for simplicity. In this case, we can determine the intermolecular potential semi-empirically. The empirical interaction potential, which is derived from the observed data on the transport phenomena etc., is usually spherically symmetric one such as Lennard-Jones type

$$\nu R^{-n} - \mu R^{-6}, \tag{3.6}$$

or a little more reasonable form[15)]

$$P \exp(-R/\rho) - \mu R^{-6}. \tag{3.7}$$

Then by assuming $v(r)=a/r^n$ or $a\exp(-br)$ and by comparing the expressions (3.5), after averaging over the orientation of the two molecules, with (3.6) or (3.7), we can easily determine the best values of the parameters. The van der Waals force potential, on the other hand, can be estimated employing the empirical or the theoretical values of the permanent quadrupole moment (the dipole moment is zero for homonuclear case), the polarizability and its anisotropy etc., for both molecules, and this potential is given by

> (*interactions between permanent multipoles*)
> + (*interactions between permanent multipoles and induced multipoles*)
> + (*interactions between induced multipoles*). $\qquad$ (3.8)

The last term, e.g., is usually given in a form: const. $\alpha_1\alpha_2/R^6$, α_1, α_2 being the polarizabilities of two molecules. But the constant coefficient is not certain from the simple theoretical calculations, and thus it will be better to adjust such coefficient by comparing (3.8) with the attractive part of the empirical potential (3.6) or (3.7).

In this way, the dependence of $V'=V_r+V_a$ on the orientations of the two molecules will be obtained. But for the calculation of the vibrational transition probability it

is also required to determine the dependence on the internuclear distances $\bar{\xi}_1, \bar{\xi}_2$. For this sake, we must know the dependence of permanent multipoles and polarizabilities on $\bar{\xi}_1$ and $\bar{\xi}_2$. For the polarizabilities, the intensity of the Raman spectra will serve for getting the required knowledge. For the permanent multipoles, however, the theoretical estimations are needed. Of course, the rigorous calculations of the multipoles of the molecule are very hard to do and the results are not reliable enough in the absolute scale. But it may be said that the ratio of the variation of the multipole to the multipole itself is reliable, though there is no authentic foundation for this expectation, at present.

For the heteronuclear molecules, the situation becomes more complex, because of the presence of the dipole moment and many undetermined parameters in V_r. But even in this case, the experimental value will serve to renormalize the uncertain results of the pure theoretical calculations.

Finally, we shall consider briefly the symmetry property of the intermolecular potential V'' for the pair of the identical molecules. For this sake, it is convenient to adopt the expression of V' in the same form as $(3 \cdot 3)$, i.e.

$$V' = \sum_{\alpha\beta\gamma} v_{\alpha\beta\gamma}'(\bar{\xi}_1, \bar{\xi}_2, R) P_\alpha(\cos \chi_1) P_\beta(\cos \chi_2) P_\gamma(\cos \chi'). \qquad (3 \cdot 9)$$

First, we shall neglect the vibration and put $\bar{\xi}_1 = \bar{\xi}_2$. Then the potential becomes symmetric for the exchange of χ_1 and $\pi - \chi_2$ for the fixed value of $|\omega_1 - \omega_2|$. (Atoms 1 and 3, 2 and 4 are assumed to be identical, resp.) Then it can be shown (Appendix I) that there are the following relations between the coefficients in the expression $(3 \cdot 9)$

$$v_{\alpha\beta\gamma}' = (-1)^{\alpha+\beta} v_{\beta\alpha\gamma}'. \qquad (3 \cdot 10)$$

Therefore, if we divide the potential V'' into two parts:

$$V' = V^+ + V^- \;\; ; \;\; V^+ = \sum_{\alpha+\beta=\text{even}} \cdots, \;\; V^- = \sum_{\alpha+\beta=\text{odd}} \cdots, \qquad (3 \cdot 11)$$

it is easy to see that V^+ is a symmetric function for the interchange of the pairs of coordinates, (θ_1, φ_1) and (θ_2, φ_2), and V^- is an antisymmetric one. Especially, for homonuclear cases the simultaneous substitutions

$$(\chi_1, \omega_1) \rightarrow (\pi - \chi_1, \pi + \omega_1), \quad (\chi_2, \omega_2) \rightarrow (\pi - \chi_2, \pi + \omega_2)$$

do not change the potential function, because by these substitutions the geometrical configuration is unaltered. From this fact, it results that there is no antisymmetric part V^- in these cases. Thus the potential between two identical homonuclear diatomic molecules is symmetric for the interchange of (θ_1, φ_1) and (θ_2, φ_2).

If we take the vibrational degrees of freedom into account, we have

$$v_{\alpha\beta\gamma}'(\bar{\xi}_1, \bar{\xi}_2, R) = (-1)^{\alpha+\beta} v_{\beta\alpha\gamma}'(\bar{\xi}_2, \bar{\xi}_1, R). \qquad (3 \cdot 12)$$

For the case of homonuclear molecules, we have $V^- = 0$ again.

§ 4.   Effect of statistics for the identical pair of molecules

For the collision between the identical molecules, we must use the properly symmetrized

wave function. Thus, instead of

$$\zeta_N(X_1)\zeta_{N'}(X_2)\varphi(\boldsymbol{R}),$$

we must employ

$$(1/\sqrt{2})[\zeta_N(X_1)\zeta_{N'}(X_2)\varphi(\boldsymbol{R}) \pm \zeta_N(X_2)\zeta_{N'}(X_1)\varphi(-\boldsymbol{R})]$$

$$= (1/\sqrt{2})[\zeta_N(X_1)\zeta_{N'}(X_2) \pm \zeta_N(X_2)\zeta_{N'}(X_1)]\cdot(1/2)[\varphi(\boldsymbol{R})+\varphi(-\boldsymbol{R})]$$

$$+ (1/\sqrt{2})[\zeta_N(X_1)\zeta_{N'}(X_2) \mp \zeta_N(X_2)\zeta_{N'}(X_1)]\cdot(1/2)[\varphi(\boldsymbol{R})-\varphi(-\boldsymbol{R})],$$

where $\zeta_N(X)$ means the internal wave function of a molecule, i.e., abbreviation of $u_{nl}(\xi)Y_{lm}(\theta,\varphi)/\xi$ and N, X stand for (n,l,m) and (ξ,θ,φ) respectively. The upper of the double sign corresponds to the Bose statistics and the lower to the Fermi statistics. In the following, however we shall confine ourselves to the Bose statistics only. Of course, the treatment for the Fermi statistics can be obtained similary.

Now we shall define the symmetrized basis

$$Z_{NN'}^{\pm} = (1/\sqrt{2})[\zeta_N(X_1)\zeta_{N'}(X_2) \pm \zeta_N(X_2)\zeta_{N'}(X_1)], \quad (N \neq N') \quad (4\cdot1)$$

$$Z_{NN}^{+} = \zeta_N(X)\zeta_N(X), \qquad Z_{NN}^{-} = 0,$$

then the wave function of the system may be written as follows

$$\Psi = \sum_{NN'}[Z_{NN'}^{+}\varphi_{NN'}^{+}(\boldsymbol{R}) + Z_{NN'}^{-}\varphi_{NN'}^{-}(\boldsymbol{R})], \quad (4\cdot2)$$

where φ^+ and φ^- mean the functions which are symmetric and antisymmetric for the substitution $\boldsymbol{R}\to-\boldsymbol{R}$ respectively. Inserting this expression into the Schroedinger equation, we have the following set of equations

$$(\nabla^2+k_{NN'}^2)\varphi_{NN'}^{+}(\boldsymbol{R}) = 2M\sum_{PP'}(Z_{NN'}^{+}|V^+|Z_{PP'}^{+})\varphi_{PP'}^{+}(\boldsymbol{R})$$

$$+ 2M\sum_{PP'}(Z_{NN'}^{+}|V^-|Z_{PP'}^{-})\varphi_{PP'}^{-}(\boldsymbol{R}),$$

$$(\nabla^2+k_{NN'}^2)\varphi_{NN'}^{-}(\boldsymbol{R}) = 2M\sum_{PP'}(Z_{NN'}^{-}|V^+|Z_{PP'}^{-})\varphi_{PP'}^{-}(\boldsymbol{R}) \qquad (4\cdot3)$$

$$+ 2M\sum_{PP'}(Z_{NN'}^{-}|V^-|Z_{PP'}^{+})\varphi_{PP'}^{+}(\boldsymbol{R}),$$

where V^+ and V^- are the parts of the potential V which are symmetric and antisymmetric for the exchange of X_1 and X_2 as defined in the last section.

φ^+ and φ^- can further be expanded as follows

$$\varphi_{NN'}^{+}(\boldsymbol{R}) = \sum_{j=\text{even}}\sum_{\nu}(2j+1)^{1/2}\{G_{NN'}^{j\nu}(R)/R\}Y_{j\nu}(\theta\varPhi) \qquad (4\cdot4)$$

$$\varphi_{NN'}^{-}(\boldsymbol{R}) = \sum_{j=\text{odd}}\sum_{\nu}(2j+1)^{1/2}\{G_{NN'}^{j\nu}(R)/R\}Y_{j\nu}(\theta\varPhi),$$

and inserting these expansions into the equation $(4\cdot3)$, we get

$$\{d^2/dR^2+k_{NN'}^2-j(j+1)/R^2\}G_{NN'}^{j\nu}$$

$$= 2M\sum_{j'=\text{even}}\sum_{\nu'PP'}(2j'+1/2j+1)^{1/2}(Y_{j\nu},Z_{NN'}^{+}|V^+|Y_{j'\nu'},Z_{PP'}^{+})G_{PP'}^{j'\nu'} \qquad (4\cdot5)$$

$$+ 2M\sum_{j'=\text{odd}}\sum_{\nu'PP'}(2j'+1/2j+1)^{1/2}(Y_{j\nu},Z_{NN'}^{+}|V^-|Y_{j'\nu'},Z_{PP'}^{-})G_{PP'}^{j'\nu'}.$$

for even j, and similar set of equations can be written down easily for odd values of j.

If the non-diagonal elements of the potential function are assumed to be small, we can neglect them in the treatment of the elastic scattering. Then, for the case $N=N'$, we have

$$\Psi = Z_{NN}^+ \varphi_{NN}^+$$

and

$$(\nabla^2 + k_{NN}^2)\varphi_{NN}^+ = 2M(Z_{NN}^+|V^+|Z_{NN}^+)\varphi_{NN}^+,$$

thus the orbital angular momentum j is restricted to be even integer. For $N \neq N'$, the total wave function becomes

$$\Psi = Z_{NN'}^+ \varphi_{NN'}^+ + Z_{NN'}^- \varphi_{NN'}^-,$$

and the set of equations becomes

$$(\nabla^2 + k_{NN'}^2)\varphi_{NN'}^+ = 2M(Z_{NN'}^+|V^+|Z_{NN'}^+)\varphi_{NN'}^+ + 2M(Z_{NN'}^+|V^-|Z_{NN'}^-)\varphi_{NN'}^-$$

$$(\nabla^2 + k_{NN'}^2)\varphi_{NN'}^- = 2M(Z_{NN'}^-|V^+|Z_{NN'}^-)\varphi_{NN'}^- + 2M(Z_{NN'}^-|V^-|Z_{NN'}^+)\varphi_{NN'}^+.$$

For the homonuclear diatomic molecules, these equations are separated as

$$(\nabla^2 + k_{NN'}^2)\varphi_{NN'}^+ = 2M\{(NN'|V^+|NN') + (NN'|V^+|N'N)\}\varphi_{NN'}^+$$

$$(\nabla^2 + k_{NN'}^2)\varphi_{NN'}^- = 2M\{(NN'|V^+|NN') - (NN'|V^+|N'N)\}\varphi_{NN'}^-,$$

where

$$(NN'|V^+|PP') \equiv \int \zeta_N^*(X_1)\zeta_{N'}^*(X_2)\, V^+ \zeta_P(X_1)\zeta_{P'}(X_2)dX_1 dX_2.$$

It may be noted that the exchange potential $(NN'|V^+|N'N)$ appears in this case which does not appear in the unsymmetrized treatment. But, these exchange potential may be as small as the other non-diagonal terms which have been neglected in our treatment.

Next, we shall consider the inelastic collision. Here again we shall assume $V^-=0$ for simplicity, and compare the two processes

$$(N) + (N) \rightleftarrows (N) + (N') \tag{4·6a} \tag{4·6b}$$

$$(N'') + (N) \rightleftarrows (N'') + (N'), \tag{4·7a} \tag{4·7b}$$

where (N) stands for the molecule in the N-state. In the unsymmetrized theory, the cross sections of these processes will be proportional to the following square of matrix elements

$$Q_{(4·6a)} \propto |(NN'|V^+|NN)|^2$$

$$Q_{(4·7a)} \propto |(N''N'|V^+|N''N)|^2$$

and $Q_{(4·6a)} \simeq Q_{(4·7a)}$ in the first approximation. Here and in the following, the equality $\simeq$ corresponds to the linear approximation of the perturbation potential, i.e.

$$(NN'|V^+|PP')$$

$$\simeq a\delta(N, P)\delta(N', P') + b\delta(N, P) \cdot (N'|v'|P') + c\delta(N', P')(N|v|P).$$

According to the Bose statistics, on the other hand, we have

$$Q_{(4 \cdot 6a)} \propto |(Z_{NN'}^+|V^+|Z_{NN}^+)|^2 = 2|(NN'|V^+|NN)|^2$$

$$Q_{(4 \cdot 7a)} \propto |(Z_{N''N'}^{\pm}|V^+|Z_{N''N'}^{\pm})|^2 \simeq |(N''N'|V^+|N''N)|^2,$$

and in the first approximation $Q_{(6a)} \simeq 2Q_{(7a)}$.

Now, in the unsymmetrized treatment of $(4 \cdot 6a)$, we must take into account the possibility that both molecule can suffer the transition. Thus the effective cross section of inelastic collision is obtained from the above value by multiplying 2. It is not the case in the Bose statistics, because the only one final state $Z_{NN'}^+$ is permissible for this case. As a result, approximate relation $Q_{(6a)} \simeq 2Q_{(7a)}$ is valid for both unsymmetrized theory and the symmetrized treatment. Furthermore, the collision number in gas phase is reduced to half for the identical pair of molecules. Thus the reaction rates of $(4 \cdot 6a)$ and $(4 \cdot 7a)$ are approximately the same. From the principle of detailed balance, this equality must hold for the inverse processes $(4 \cdot 6b)$ and $(4 \cdot 7b)$ too. It can be confirmed easily. For the unsymmetrized theory, we have

$$Q_{(4 \cdot 6b)} \propto |(NN|V^+|NN')|^2$$

$$Q_{(4 \cdot 7b)} \propto |(N''N|V^+|N''N')|^2,$$

and in the first approximation $Q_{(6b)} \simeq Q_{(7b)}$. In the symmetrized treatment, on the other hand, we have

$$Q_{(4 \cdot 6b)} \propto \tfrac{1}{2}|(Z_{NN}^+|V^+|Z_{NN'}^+)|^2 = |(NN|V^+|NN')|^2$$

$$Q_{(4 \cdot 7b)} \propto |(Z_{N''N}^{\pm}|V^+|Z_{N''N'}^{\pm})|^2 \simeq |(N''N|V^+|N''N')|^2,$$

where the factor $\tfrac{1}{2}$ corresponds to the fact that only the partial waves with even orbital angular momenta j can be combined with the final state. Thus we have $Q_{(6b)} \simeq Q_{(7b)}$. Now this time the collision numbers are equal and, therefore, the reaction rates are equal for $(4 \cdot 6b)$ and $(4 \cdot 7b)$ as expected. (On these effects of the statistics, our previous considerations have been mistaken, in which we concluded $Q_{(6b)} \simeq \tfrac{1}{2}Q_{(7b)}$.[14])

In this section we have employed the (j, ν, l, m, l', m')-representation. However, this can be transformed easily into the (J, j, L, l, l', μ)-representation.

§ 5.   General discussions on the cross section

In this section, the general asymptotic form of the scattered wave function is considered and then, in relation to this asymptotic form, we proceed to the definition of the cross sections of the elastic and inelastic scatterings. As is well known the plane wave is expanded as follows

$$\exp(ikR\cos\Theta) = \sum_{j=0}^{\infty}(2j+1)i^j P_j(\cos\Theta)\bar{f}_j(R),$$

where $\bar{f}$ behaves asymptotically as

$$\bar{f}_j(R) \sim \sin(kR - \tfrac{1}{2}j\pi)/kR \qquad \text{for } R \to \infty.$$

Thus the incident wave for our problem can be written, except for the vibrational part, as follows

$$\Psi_0 = Y_{lm}(\theta_1\varphi_1)\, Y_{l'm'}(\theta_2\varphi_2)\exp(ikR\cos\theta)$$

$$= \sum_{j=0}^{\infty}(2j+1)i^j\,\bar{f}_j(R)\,(4\pi/2j+1)^{1/2}\,Y_{j0}(\theta\Phi)\,Y_{lm}(\theta_1\varphi_1)\,Y_{l'm'}(\theta_2\varphi_2) \qquad (5\cdot1)$$

$$= (1/R)\sum_{j=0}^{\infty}\sum_{J,L}\bar{F}_{JjL}^{lml'm'}\,\mathbf{Y}_{jLll'}^{Jm+m'},$$

where $\bar{F}$ behaves asymptotically as

$$\bar{F}_{JjL}^{lml'm'}\sim s_{J0\;m+m'}^{jL}\,s_{Lmm'}^{ll'}\,[4\pi(2j+1)]^{1/2}i^j\sin(kR-\tfrac{1}{2}j\pi)/k$$

$$= s_{J0\;m+m'}^{jL}\,s_{Lmm'}^{ll'}\,[\pi(2j+1)]^{1/2}\{\exp(ikR)-(-1)^j\exp(-ikR)\}/ki, \qquad (R\to\infty).$$

The vibrational part can be taken into account easily. But in the thermal collisions between diatomic molecules, the vibrational transition takes place with very small probability and its influence on the elastic scattering or the rotational transition is negligible. For this reason, we shall neglect the vibrational degrees of freedom in this section. Now we can assume the following form for the elastically scattered wave

$$\Psi_{el} = (1/R)\sum_{j=0}^{\infty}\sum_{J,L}E_{JjL}^{lml'm'}\,\mathbf{Y}_{jLll'}^{Jm+m'} \qquad\qquad (5\cdot2)$$

$$E_{JjL}^{lml'm'}\sim s_{J0\;m+m'}^{jL}\,s_{Lmm'}^{ll'}\,[\pi(2j+1)]^{1/2}(S_{JjL}^{lml'm'}-1)\exp(ikR)/ki, \qquad (R\to\infty).$$

The inelastically scattered wave, on the other hand, may be written as

$$\Psi_{in} = (1/R)\sum_{\lambda\lambda'}\sum_{j=0}^{\infty}\sum_{J,L}G_{JjL\lambda\lambda'}^{lml'm'}\,\mathbf{Y}_{jL\lambda\lambda'}^{Jm+m'} \qquad (\lambda,\lambda')\neq(l,l') \qquad (5\cdot3)$$

$$G_{JjL\lambda\lambda'}^{lml'm'}\sim s_{J0\;m+m'}^{jL}\,s_{Lmn'}^{ll'}\,[\pi(2j+1)]^{1/2}T_{JjL\lambda\lambda'}^{lml'm'}\exp(ik_f R)/k_f i, \qquad (R\to\infty).$$

k_f is the wave number corresponding to the final state, which is defined by $(2\cdot6)$. In the asymptotic form $(5\cdot3)$ the factors before T can be, of course, included in the undetermined factor T. But the above form is adopted to make equal its form with the case of elastic scattering.

Now the cross section can be expressed easily by making use of the scattering matrix S. By imagining a large spherical surface (SS) with the radius R_0, the outgoing flux of the wave Ψ_{el} across this SS is easily calculated as

$$I_{el} = (1/2iM)\int_{SS}R_0^2\,d\Omega\iint d\Omega_1\,d\Omega_2(\Psi_{el}^*\cdot\partial\Psi_{el}/\partial R-\partial\Psi_{el}^*/\partial R\cdot\Psi_{el})$$

$$= (k/M)(\pi/k^2)\sum_j\sum_{J,L}(2j+1)(s_{J0\;m+m'}^{jL}\,s_{Lmm'}^{ll'})^2|S_{JjL}^{lml'm'}-1|^2,$$

where we have used the orthonormality of $\mathbf{Y}$. The flux of the incident plane wave, on the other hand, is clearly

$$I_0 = v = k/M \quad\text{per unit area.}$$

Thus the cross section for the elastic scattering can be defined as usual

$$Q_{el} \equiv I_{el}/I_0 = (\pi/k^2)\sum_j (2j+1)\sum_{J,L}(s^{jL}_{J0\,m+m'}\,s^{ll'}_{Lmm'})^2\,|S^{lml'm'}_{JjL}-1|^2. \qquad (5\cdot4)$$

Now the continuity equation must hold for the total wave function $\Psi = \Psi_0 + \Psi_{el} + \Psi_{in}$, which means that the net incoming flux of $\Psi_0 + \Psi_{el}$ must be equal to the flux of the outgoing wave Ψ_{in}. (Remember that $\Psi_0 + \Psi_{el}$ and Ψ_{in} are orthogonal to each other because of different internal wave functions.) From this fact, the cross section for the inelastic scattering can be given formally by the integral

$$Q_{in} = -\frac{1}{2iM}\int_{ss} R_0^2\,d\Omega \iint d\Omega_1\,d\Omega_2 \Big\{(\Psi_0^* + \Psi_{el}^*)\frac{\partial}{\partial R}(\Psi_0 + \Psi_{el})$$
$$-\frac{\partial}{\partial R}(\Psi_0^* + \Psi_{el}^*)\cdot(\Psi_0 + \Psi_{el})\Big\}\Big/I_0.$$

In calculating the integral, the terms which contain Ψ_0 and Ψ_0^* only cancel with each other, and the terms which contain Ψ_{el} and Ψ_{el}^* only become $-Q_{el}$. Thus we have

$$Q_{in} = -Q_{el} + \frac{1}{2ik}\int_{ss} R_0^2\,d\Omega \iint d\Omega_1\,d\Omega_2 \Big(\frac{\partial \Psi_0^*}{\partial R}\Psi_{el} + \frac{\partial \Psi_{el}^*}{\partial R}\Psi_0 - \Psi_0^*\frac{\partial \Psi_{el}}{\partial R} - \Psi_{el}^*\frac{\partial \Psi_0}{\partial R}\Big)$$
$$= (\pi/k^2)\sum_j(2j+1)\sum_{J,L}(s^{jL}_{J0\,m+m'}\,s^{ll'}_{Lmm'})^2\{1-|S^{lml'm'}_{JjL}|^2\}. \qquad (5\cdot5)$$

It can be seen from this, that the cross section for the inelastic scattering is determined by the scattering matrix S which determines the elastic scattering. Properly speaking, however, matrix S itself can be determined by taking not only the elastic but also the inelastic processes into account. Of course, the item of the inelastic processes can not be expressed by S, but by T. Using the matrix T, the particular inelastic collision cross section $l \to \lambda$, $l' \to \lambda'$ can be given by

$$Q_{in} = (\pi/kk_f)\sum_j(2j+1)\sum_{J,L}(s^{jL}_{J0\,m+m'}\,s^{ll'}_{Lmm'})^2\,|T^{lml'm'}_{JjL\lambda\lambda'}|^2.$$

Summing up such cross sections over all possible inelastic processes and equating the result with $(5\cdot5)$, we have the following relation between S and T,

$$\sum_j(2j+1)\sum_{J,L}(s^{jL}_{J0\,m+m'}\,s^{ll'}_{Lmm'})^2\Big\{1-|S^{lml'm'}_{JjL}|^2 - \sum_{\lambda\lambda'}\frac{k}{k_f}|T^{lml'm'}_{JjL\lambda\lambda'}|^2\Big\}=0.$$

Even if the inelastic processes are forbidden energetically or take place with very small probabilities, the absolute value of the elements of matrix S are not equal to unity in general. That is, even if the expression $(5\cdot5)$ vanishes, some of the $|S^{lml'm'}_{JjL}|$ may be larger and some others smaller than unity, corresponding to the possibility of the change of j and L during the collision. Such redistribution in j and L can take place easily for the non-spherical scattering potential. For this reason, it is impossible to express the elastic scattering by the phase shifts only, as in the spherical fields. However, if we can assume in the first approximation that the quantum numbers j, l, l', L are good quantum numbers (J, $m+m'$ are good quantum numbers, rigorously.), the absolute values of matrix elements of S become unity and, therefore, by putting

$$S^{lml'm'}_{.JjL} = \exp(2i\,\partial^{JL}_{jll'})$$

we have immediately that

$$Q_{el} = (4\pi/k^2)\sum_{j}(2j+1)\sum_{J,L}(s^{jL}_{J0\,m+m'}\,s^{ll'}_{L.mm'})^2\sin^2\partial^{JL}_{jll'}\,,\qquad(5\cdot6)$$

which is also obtained directly by replacing the $\sin^2\partial_j$ in the well-known formula

$$Q_{el} = (4\pi/k^2)\sum_{j}(2j+1)\sin^2\partial_j\quad(central\ field)$$

by its weighted mean

$$\sum_{J,L}(s^{jL}_{J0\,m+m'}\,s^{ll'}_{L.mm'})^2\sin^2\partial^{JL}_{jll'}\,.$$

The cross sections which are effective for the viscosity or the diffusion can be obtained straightforwardly, but these have not such simple forms. For instance, for the case of negligible inelastic scattering, the cross section for the viscosity is given formally by

$$Q_{vis} = \iiint d\Omega\,d\Omega_1\,d\Omega_2\sin^2\Theta\,\Big|\sum_{jJL}F^{lml'm'}_{JjL}\,Y^{Jm+m'}_{jLll'}\Big|^2.$$

But this integral does not become simple, even if we assume that the absolute value of the matrix elements of S are unity. (See Appendix II.)

§ 6. More detailed discussions on the elastic scattering

As in the previous section, we shall neglect the vibrational degrees of freedom in this section too. It is easy to see by analyzing the experimental data of ultrasonic dispersion that the probabilities of the vibrational transitions are very small under the ordinary conditions.[16] For instance, the deactivating transition of an excited O_2 molecule takes place only once in 160000 collisions in the O_2 gas for the room temperature. For an excited Cl_2 molecule in Cl_2 gas, transition takes place once in 34000 collisions. For Cl_2 in HCl gas, which is one of the most favorable cases, once in 120 collisions. The probabilities of exciting processes to the upper vibrational states are much smaller than these figures. Thus we may safely neglect the vibrational transitions in the first approximation.

Now the most simple method of treatment of the elastic scattering is to reduce the problem to the scattering under the spherically symmetric potential field which is obtained from the non-spherical potential by averaging over the orientations of two molecules. In fact, even if the two colliding molecules have the definite rotational angular momenta l, l' respectively, the z-components of these angular momenta will change from one particular collision to another and, therefore, the effective scattering process must be obtained by averaging over the quantum numbers m, m' which define the z-components of the angular momenta. The average of l'' over m and m' leads, indeed, to the spherical potential which is independent of l and l'. Property speaking, however, the averaging must be made not for the potential, but for the cross section, and now we shall investigate the difference of these two averaging procedures, by assuming that l, l', L j (and of course J) are good quantum numbers.

We denote the averaged potential field by V_0, and the phase shifts of the scattered

wave from this field V_0 by ∂_j. Then the scattering cross section for this potential is given by

$$Q_{el}^0 = (4\pi/k^2)\sum_j (2j+1)\sin^2\partial_j , \qquad (6\cdot1)$$

k being the wave number. Next, we shall denote the phase shifts for the actual potential field[*]

$$(J, j, L, l, l' \,|V'|\, J, j, L, l, l')$$

by

$$\partial_{jll'}^{JL} = \partial_j + \varDelta\partial_{jll'}^{JL} ,$$

and the cross section calculated by these phase shifts by

$$Q_{el} = Q_{el}^0 + \varDelta Q_{el} .$$

Q_{el}^0 is a function of the wave number k only, but Q_{el} depends on l, l', m, m' too. If $\varDelta\partial$'s are assumed to be small, we can get from the equation $(5\cdot6)$ that

$$\varDelta Q_{el} = (4\pi/k^2)\sum_j (2j+1)\sin 2\partial_j \sum_{J,L} (s_{J0\,m+m'}^{jL}\, s_{Lmm'}^{ll'})^2 \varDelta\partial_{jll'}^{JL} \qquad (6\cdot2)$$

in the first approximation. The change of the phase shifts appeared in this formula can be given by the integral

$$\varDelta\partial_{jll'}^{JL} \doteqdot (1/k)\int_0^\infty dR |F_j(R)|^2 \varDelta V_{jll'}^{JL} \qquad (6\cdot3)$$

in the first approximation, where $\varDelta V_{jll'}^{JL}$ means the potential change

$$\varDelta V_{jll'}^{JL} = (J, j, L, l, l' \,|V'|\, J, j, L, l, l') - V_0$$

and $F_j(R)$ is the solution of the radial equation

$$\{d^2/dR^2 + k^2 - j(j+1)/R^2 - 2MV_0(R)\}F_j(R) = 0, \qquad (6\cdot4)$$

which vanishes at the origin and is normalized as $\sin(kR - \tfrac{1}{2}j\pi + \partial_j)$ asymptotically. The formula $(6\cdot3)$ can be derived as follows. Let F_j' be the solution of the radial equation for a potential $V_0 + \varDelta V$, which is normalized as $\sin(kR - \tfrac{1}{2}j\pi + \partial_j + \varDelta\partial_j)$ asymptotically. Multiplying F_j' from the left to $(6\cdot4)$ and multiplying F_j to

$$\{d^2/dR^2 + k^2 - j(j+1)/R^2 - 2M(V_0 + \varDelta V)\}F_j'(R) = 0$$

and subtracting one from the other, we obtain

$$d/dR(F_j' \cdot kF_j/dR - F_j \cdot dF_j'/dR) = F_j \varDelta V F_j'.$$

By noting that both F_j and F_j' vanish at the origin, we get by integration

$$(F_j' \cdot dF_j/dR - F_j \cdot dF_j'/dR)_{R=\infty} = \int_0^\infty dR\, F_j \varDelta V F_j'$$

---

* Matrix elements of V' do not depend on the sixth quantum number μ, because the latter quantity define the z-component of the total angular momentum, i.e., the orientation of the system as a whole, while V' depends only on the relative configuration of the two molecules. (See also Appendix III.)

the left hand side of which can be calculated easily, resulting $k \sin \varDelta \eth_j$. This is yet a rigorous equation. Since $\varDelta \eth_j$ have been assumed to be small we may put $\sin \varDelta \eth_j \doteqdot \varDelta \eth_j$, and also $F_j' \doteqdot F_j$ which results the above equation $(6 \cdot 3)$.

As is shown in the Appendix III, we can prove that

$$\sum_{mm'JL} (s_{J0\,m+m'}^{jL} s_{L,mm'}^{lll'})^2 \varDelta V_{jll'}^{JL} = 0, \qquad (6 \cdot 5)$$

Thus, from $(6 \cdot 2)$, $(6 \cdot 3)$ and $(6 \cdot 5)$, we get immediately that

$$\sum_{mm'} \varDelta Q_{el} \doteqdot 0 \qquad (6 \cdot 6)$$

in the first approximation. Therefore, we may calculate the cross section for the spherical potential V_0, instead of the non-spherical potential V'' in the first approximation.

On the other hand, if $\varDelta V_{jll'}^{JL}$ are assumed to be small, m and m' may be considered as the good quantum numbers. Then we may calculate the cross section from the equations

$$\{d^2/dR^2 + k^2 - j(j+1)/R^2\}\, f_{j0lml'm'} = 2M\,(j0\,l\,m\,l'\,m'|V''|j0\,l\,m\,l'\,m')\, f_{j0lml'm'}$$

corresponding to $(2 \cdot 5)$. In this approximation, the cross section is given by

$$Q_{el} = (4\pi/k^2)\sum_j (2j+1)\sin^2 \delta_j^{jml'm'}$$

and we can easily verify the same conclusion as before, i.e., $\sum_{mm'} \varDelta Q_{el} \doteqdot 0$ by employing the relation

$$\sum_{mm'} \varDelta V_j^{lml'm'} = 0$$

$$\varDelta V_j^{lml'm'} \equiv (j0\,l\,m\,l'\,m'|V''|j0\,l\,m\,l'\,m') - V_0.$$

In any case, if the anisotropic part of the intermolecular potential is small, the calculation of the scattering cross section, taking the anisotropy into account in the first approximation, leads to the same result as that obtained by using the averaged potential V_0.

In this connection, it must be noted that K. Ohno and the present author have investigated recently the thermal elastic collision between H_2 molecules at low temperatures.[17] They compared the cross section of para-para collision with that of para-ortho collision and found that the difference $(6 \cdot 2)$ of the total cross sections and also the corresponding difference in the cross sections for viscosity were very small indeed, though their calculations were based on direct evaluation of the phase shifts, without employing the approximate relation $(6 \cdot 3)$.

Of course, these conclusions will be broken if the anisotropy of the intermolecular potential is large. First of all, in such a case, the quantum numbers j, L etc. can not be considered as good quantum numbers. This makes the treatment of the scattering process very difficult because we have to solve a coupled set of differential equations. Now we shall proceed to discuss such more general cases.

First, we shall consider the special case in which j is yet a good quantum number, but L is not so. Then the wave equations have the form

$$\{d^2/dR^2 + k^2 - j(j+1)/R^2\}\, F_L = 2M \sum_{L'} (L|V'|L')\, F_{L'}\,. \tag{6.7}$$

Furthermore, if V'' has the property that the matrix elements of it can be written in a form

$$(L|V'(R)|L') = (L|u|L')\, v(R),$$

in which the first factor is independent of R, then the first factor must be hermitian and can be diagonalized by suitable unitary transformation. In this case, the equations (6.7) become a set of separated equations, and, therefore, become much easier to treat. But this is a special case. In the general case, we are not in such a fortunate situation, and also j is not good quantum number.

The change of j is, of course, a kind of elastic scattering for the observer. But it is rather close to the inelastic collision processes. Probably the jump of j will take place in the neighborhood of the classically closest distance R_0, and in this region the change of j means the apparent change of the kinetic energy by an amount $j(j+1)/2MR_0^2$ in the radial equation. Thus, in order to change the j-value, the wave length of the radial wave function must change suddenly in this region, which is just the feature of the inelastic processes. Therefore, j can be a good quantum number, only if the condition

$$\textit{change of } j(j+1)/2MR_0^2 \geqslant \textit{initial kinetic energy} \tag{6.8}$$

is valid or the anisotropy of the intermolecular potential, which causes the change of j, is extremely small. Unfortunately, the condition (6.8) is not insured to be valid in the thermal collisions at room temperature, even for the most favourable case, i.e., for the hydrogen.

Now the equations to be solved are

$$\{d^2/dR^2 + k^2 - j(j+1)/R^2\}\, F_{jL}(R) = 2M \sum_{sN} (jL\cdots|V''|sN\cdots)\, F_{sN}(R)\,. \tag{6.9}$$

Since this set of equations are linear, we can obtain the required solution by superposing the particular solutions with proper amplitudes and phases. Thus, we shall first consider the problem in which the incoming wave has the definite values of j and L. In the following, the solution of the equation

$$\{d^2/dR^2 + k^2 - j(j+1)/R^2\}\, F_{jL}(R) = 2M(jL\cdots|V''|jL\cdots)\, F_{jL}(R) \tag{6.10}$$

satisfying the boundary conditions

$$F_{jL}(0) = 0, \quad F_{jL}(R) \sim \sin(kR - \tfrac{1}{2}j\pi + \delta_j) \quad (R \to \infty) \tag{6.11}$$

will be denoted by F_{jL}^0, and the deviations due to the non-diagonal matrix elements of V'' by δF_{jL}. Then we have

$$\{d^2/dR^2 + k^2 - j(j+1)/R^2 - 2M(jL|V''|jL)\}\, \delta F_{jL} = 2M \sum_{sN \neq jL} (jL|V'|sN)\, \delta F_{sN}\,,$$
$$\tag{6.12}$$

$$\{d^2/dR^2 + k^2 - s(s+1)/R^2 - 2M(sN|V''|sN)\}\, \delta F_{sN}$$
$$= 2M \sum_{pQ \neq sN} (sN|V''|pQ)\, \delta F_{pQ} + 2M\,(sN|V''|jL)\, F_{jL}^0\,.$$

Assuming that F_{jL}^0 is a quantity of zeroth order and V'' is a small quantity of the first

order, we can see that ∂F_{sN} $(sN \neq jL)$ are of the first order and ∂F_{jL} is a second order small quantity. Thus in the first approximation, $(6\cdot12)$ become

$$\partial F_{jL} = 0 ,$$

$$\{d^2/dR^2 + k^2 - s(s+1)/R^2 - 2M(sN|V'|sN)\}\partial F_{sN} = 2M(sN|V'|jL)F_{jL}^0 . \qquad (6\cdot13)$$

The first of these equations means that the damping effect is not taken into account in this approximation, while the second equation can be solved formally and its solution which represents the outgoing wave asymptotically is given by[18]

$$\partial F_{sN} = -G_{sN}(R)\int_0^R F_{sN}^0(R)(sN|V'|jL)F_{jL}^0(R)dR \qquad (6\cdot14)$$

where G_{jL} is the solution of the equation $(6\cdot10)$ which is normalized asymptotically as $(1/k)\exp\{i(kR - \frac{1}{2}j\pi + \delta_j)\}$. Thus the asymptotic form of ∂F_{sN} is

$$\partial F_{sN} \sim (1/k)\exp\{i(kR - \frac{1}{2}j\pi + \delta_s)\}\int_0^\infty F_{sN}^0(sN|V'|jL)F_{jL}^0 dR . \qquad (6\cdot15)$$

Now we shall return to our problem in which the plane wave is incoming. In this case the solution in the zeroth approximation is given by the linear combination of F_{jL}^0 with the coefficients

$$s_{J0\ m+m'}^{jL} s_{Lmm'}^{ll''} \left[4\pi(2j+1)\right]^{1/2} i^j \exp(i\delta_j)/k .$$

(See the asymptotic form of $\bar{F}$ in the previous section.) Thus the asymptotic form of the solution in the first approximation is given by

$$F_{jL}^{(J)} \sim (\sqrt{4\pi}/k)\big[s_{J0\ m+m'}^{jL} s_{Lmm'}^{ll''}(2j+1)^{1/2} i^j \exp(i\delta_j)F_{jL}^0$$

$$+ \sum_{sN} s_{J0\ m+m'}^{sN} s_{Nmm'}^{ll''}(2s+1)^{1/2} i^s \exp(i\delta_s)(1/k)\exp\{i(kR - \frac{1}{2}j\pi + \delta_j)\}$$

$$\times \int_0^\infty F_{jL}^0(jL|V'|sN)F_{sN}^0 dR\big] .$$

This solution being equated with $\bar{F}_{ljL}^{lml'm'} + F_{ljL}^{lml'm'}$ in the previous section, we have the scattering matrix

$$S_{JjL}^{lml'm'} = \exp(2i\delta_j) + (2/k)\sum_{sN}(s_{J0\ m+m'}^{sN} s_{Nmm'}^{ll''}/s_{J0\ m+m'}^{jL} s_{Lmm'}^{ll''}) \times$$

$$(2s+1/2j+1)^{1/2} i^{s-j+1} \exp(i\delta_j + i\delta_s) \times \qquad (6\cdot16)$$

$$\int_0^\infty F_{jL(ll'')}^{(J)0}(JjLll''|V'|JsNll')F_{sN(ll')}^{(J)0} dR .$$

Substitution of this expression into $(5\cdot4)$ yields the formula for Q_{el}.

The approximate solution thus obtained may be substituted in the right hand side of $(6\cdot12)$, giving the equation for the second approximation, though the solution will become more complex. If the second order correction terms thus calculated are not small enough compared with the first order correction ∂F_{jL}, this successive approximation method may be impracticable, because we must investigate whether the next, i.e. the third, approximation terms are small or not. Unfortunately, however, we have no general method to solve a set of differential equations in other way. Stueckelberg[19] reduced the coupled pair of

differential equations to a higher order equation and solve the latter by means of WKB approximation method, while Massey and Mohr[20] solved directly the coupled pair of differential equations for a simplified case. But these methods are hard to extend to the set of a number of coupled equations.

Finally, it must be noted that the non-diagonal elements which cause the inelastic collision processes may also affect the elastic collision cross section, even if the inelastic processes are forbidden energetically. In other words, some excited states may appear virtually in the intermediate stage of collision. This effect also force us to solve the set of equations in the general case, but, if this effect is not so large, we may employ the second order perturbation method. On the other hand, if this effect is large enough, the two molecules can no longer continue their free rotations and a particular relative orientation becomes favourable during the collision, which means the temporary excitation of the upper rotational states. In such a case, the method of p.s.s. (perturbed stationary state), which is familiar in the theory of atomic collisions,[21] may be more reasonable procedure of solution than the ordinary perturbation method. The essential point of this p.s.s. method is very simple. In this method, the total wave function is expanded in terms of the proper functions of

$$\left\{ -\frac{1}{2\mu_1} \varGamma_1^2 - \frac{1}{2\mu_2} \varGamma_2^2 + V_1(\xi_1) + V_2(\xi_2) + V' - E(R, \theta, \varPhi) \right\} \varPsi = 0,$$

while according to the method in the previous sections the proper functions of

$$\left\{ -\frac{1}{2M} \nabla_R^2 - \frac{1}{2\mu_1} \nabla_1^2 - \frac{1}{2\mu_2} \nabla_2^2 + V_1(\xi_1) + V_2(\xi_2) - E \right\} \varPsi = 0$$

are employed. In the former of these equations, the values of $(R, \theta, \varPhi)$ are fixed, and the kinetic energy of the relative motion of two molecules is treated as the perturbation, while in the latter method the intermolecular potential is considered as the perturbation. The general formulation of the problem according to the p.s.s. method can be given straightforwardly, but we shall not enter this problem further.

§ 7. General discussions on the inelastic collisions

The problem of the inelastic collisions is the most interesting part of our collision problem, because the inelastic processes depend essentially on the internal structures of the colliding molecules, while the main features of the elastic collisions have been obtained already from the rigid sphere model or from other spherical fields. Unfortunately, however, the straightforward calculations of the cross sections for the inelastic processes impel us to carry out the overwhelmingly hard work. Thus the simplifying assumptions must be introduced before the practical calculations.

Now the cross sections for the inelastic collisions are influenced remarkably by the following two factors. The first is the overlaping between the initial and the final translational wave functions. Applying the Born approximation for simplicity, the transition probability will proportional to the absolute square of the matrix element

$$V_{if} = \int \exp\{-i(\mathbf{k}_f - \mathbf{k}_i)\mathbf{R}\} \cdot (\alpha_f|V''(\mathbf{R})|\alpha_i)d\mathbf{R}$$

where α_i, α_f indicate the internal state of the molecules; $\mathbf{k}_i$, $\mathbf{k}_f$ the wave numbers in the initial and the final states respectively. In the thermal collisions, the wave length of the relative motion of the two molecules is usually less than 1 A.U., while the matrix element $(\alpha_f|V''|\alpha_i)$ changes more slowly as a function of R. Thus, V_{if} will be small in absolute value, unless $\Delta k = k_f - k_i$ is small. Of course, the Born approximation is not adequate for the molecular collisions in the thermal energy region, but the essential feature must not be changed when the plane wave $\exp(i\mathbf{k}\mathbf{R})$ is replaced by the suitable distorted wave. Thus, we may conclude that if the difference between the wave number of the initial translational motion and that of the final one is very large, the transition probability will be very small.

It is easy to see that the inelastic collision can hardly occur in the low energy collisions. For endothermic reactions, transition can not take place, of course, if the initial energy is less than the threshold energy. If the initial energy is just above the threshold one, the change in the wave number will be large. And even for the exothermic reactions, the wave number will change appreciably if the initial kinetic energy is not large enough. In another words, in order to get the large probability for transition, the initial kinetic energy must be large enough compared with the energy change in the inelastic process. On the other hand, if the initial velocity is extremely large, there will be no time for the transition to take place during the collision and, therefore, the cross section will be small. The cross section may become large for a intermediate region between these two extreme cases.

The second factor which influences the inelastic collision cross section is the dependence of the intermolecular force potential on the coordinate in which the transition takes place. Thus in order that the vibrational transition takes place with an appreciable probability, a small change in the internuclear distance of a molecule must bring the considerable change in the intermolecular force and in order to have the rotational transition it is necessary that the molecular potential deviates very much from the spherical symmetry. Now the intermolecular potential which is effective for the thermal collisions is determined by the electron distributions in the outermost shells of both molecules. It is clear that in a molecule the force field, under which the outermost electron moves, is something like the Coulomb field of the hydrogen nucleus which is placed in the center of the molecule, especially when the electron considered is far apart from the center. It means that the molecular orbital for the outermost electron will become something like hydrogen-atomic-orbital asymptotically and does not depend very much on the nuclear distance and on the orientation of the molecular axes. For this reason, the intermolecular force becomes rapidly spherically symmetric for the larger values of the intermolecular separation, and also becomes independent of the internuclear distances in the two molecules. Therefore, in order that the inelastic process takes place, two molecules must approach closely to each other and also for this sake the incident energy must be large.

Now the vibrational level spacings in the ordinary diatomic molecules are much larger

than the energy of the thermal motion in the room temperature. Thus, it is expected from the above considerations, that the transition probability is very small under the ordinary thermal condition. This expectation has been confirmed by the experimental data mentioned in the beginning of the last section. The rotational transition in the hydrogen molecule also takes place with very small probability because of its rather large level spacing.[*] Thus for the vibrational transitions in the ordinary diatomic molecules and for the rotational transitions in some special molecules such as H_2, the first order approximation in the perturbation method i.e. the method of distorted wave may be applied to calculate the cross sections, while for the rotational transitions in the other molecules we must solve a coupled set of equations because the transition probabilities are rather large for these cases. Of course, the straightforward solution of such set of differential equations is very difficult as was seen in the elastic scattering problem and we must simplify the problem as far as possible before solving it. For these cases with small level spacings, the translational energy does not change appreciably in the transition process. Furthermore, the motion of a molecule as a whole is very similar to that of the classical particle, having a rather definite orbit with a definite velocity, because of its large mass. Therefore, the translational motion may be treated approximately in the classical manner, i.e., we may determine the translational motion by solving the classical equation of motion under the averaged potential field. Then the relative distance of the two molecules is given by a function of the time. The transition will be caused by the time-dependent perturbation. Such semi-classical treatments of the inelastic collision processes have been reported frequently by many authors.[24] These authors have investigated only the first order perturbation problem by this method. But the strong coupling case may also be treated in this way, if the kinetic energy does not change appreciably during the collision.

§ 8. Rotational transitions

Now we shall consider the rotational transitions disregarding the vibrational degrees of freedom. The special case of H_2 can be treated by the distorted wave approximation in exactly the same way as in the vibrational transitions which are discussed in the next section. In this paper, therefore, we shall confine ourselves to mention that the problem of H_2 has been investigated by Roy and Rose previously[11] and a little more thorough calculations have also been made recently by the present author.[14]

In the following we shall consider the rotational transition in which the energy exchange between the translation and the rotation is rather small. We shall employ the semi-classical method, as discussed above. For this sake, we must start from the time-dependent Schroedinger equation:

$$i\, \partial \Psi / \partial t = \{(-1/2M)\nabla_R^2 + H_0 + V'\}\, \Psi , \qquad (8\cdot 1)$$

---

* It is well known that the transition between para- and ortho-hydrogens is usually forbidden. Thus the rotational transition between the neighboring levels is forbidden in the hydrogen molecule. This fact, together with the small moment of intertia, makes the effective level spacings very large. The experimental evidence for the small transition probability in H_2 have been given by Stewart[22] and Rhodes.[23]

where H_0 means the Hamiltonian operator for the internal degrees of freedom of the two colliding molecules. The wave function Ψ may be expanded as

$$\Psi(\boldsymbol{R}, X_1, X_2, t) = \sum_{\alpha'\alpha''} \zeta_{\alpha'}(X_1)\zeta_{\alpha''}(X_2)(2\pi)^{-3/2}$$

$$\times \int f_{\alpha'\alpha''}(\boldsymbol{k}, t)\exp\{i\boldsymbol{k}\boldsymbol{R} - i(E_k + E_{\alpha'} + E_{\alpha''})t\}d\boldsymbol{k} \qquad (8\cdot2)$$

or

$$= \sum_{\alpha} \zeta_{\alpha}(X)(2\pi)^{-3/2}\int f_{\alpha}(\boldsymbol{k}, t)\exp\{i\boldsymbol{k}\boldsymbol{R} - i(E_k + E_{\alpha})t\}d\boldsymbol{k} \qquad (8\cdot2\text{a})$$

for brevity, where $\zeta_{\alpha'}(X_1)\zeta_{\alpha''}(X_2)$ (or $\zeta_\alpha(X)$) stand for the orthonormalized internal wave functions and

$$E_k = k^2/2M, \quad E_\alpha = internal\ energy\ in\ the\ state\ \alpha.$$

By inserting $(8\cdot2\text{a})$ into $(8\cdot1)$, and taking the equation

$$\{(-1/2M)\Gamma_R^2 + H_0\}\exp(i\boldsymbol{k}\boldsymbol{R})\zeta_\alpha(X) = (E_k + E_\alpha)\exp(i\boldsymbol{k}\boldsymbol{R})\zeta_\alpha(X) \qquad (8\cdot3)$$

into account, we get

$$i\sum_{\alpha}\zeta_\alpha(X)\int\frac{\partial f_\alpha}{\partial t}\exp\{i\boldsymbol{k}\boldsymbol{R} - i(E_k + E_\alpha)t\}d\boldsymbol{k}$$

$$= V''\sum_{\alpha'}\zeta_{\alpha'}(X)\int f_{\alpha'}\exp\{i\boldsymbol{k}'\boldsymbol{R} - i(E_{k'} + E_{\alpha'})t\}d\boldsymbol{k}'$$

which gives, by the orthonormality of $\zeta_\alpha(X)$,

$$i\int\frac{\partial f_\alpha}{\partial t}\exp(i\boldsymbol{k}\boldsymbol{R} - iE_k t)d\boldsymbol{k} = \sum_{\alpha'}(\alpha|V''|\alpha')\int f_{\alpha'}\exp\{i\boldsymbol{k}'\boldsymbol{R} - iE_{k'}t + i\varDelta E_{\alpha\alpha'}t\}d\boldsymbol{k}'$$

where

$$(8\cdot4)$$

$$(\alpha|V''|\alpha') = \int\zeta_\alpha^*(X)V''\zeta_{\alpha'}(X)dX, \quad \varDelta E_{\alpha\alpha'} = E_\alpha - E_{\alpha'}.$$

Now the function

$$\varphi_\alpha(\boldsymbol{R}, t) = (2\pi)^{-3/2}\int f_\alpha(\boldsymbol{k}, t)\exp(i\boldsymbol{k}\boldsymbol{R} - iE_k t)d\boldsymbol{k} \qquad (8\cdot5)$$

represents the wave packet and we may assume that this function takes the appreciable values only in the limited region in space for each time t.

Here we shall factorize the function f_α in the following way:

$$f_\alpha = C_\alpha \cdot f_\alpha^0, \qquad (8\cdot6)$$

$$\int|f_\alpha^0|^2 d\boldsymbol{k} = 1, \qquad (8\cdot6\text{a})$$

$$i\int\frac{\partial f_\alpha^0}{\partial t}\exp(i\boldsymbol{k}\boldsymbol{R} - iE_k t)d\boldsymbol{k} = (\alpha|V''|\alpha)\int f_\alpha^0\exp(i\boldsymbol{k}'\boldsymbol{R} - iE_{k'}t)d\boldsymbol{k}' \qquad (8\cdot7\text{b})$$

i.e., f_α^0 is a solution representing the elastic scattering. Generally speaking, $C_\alpha \cdot f_\alpha^0$ cannot be a solution of $(8\cdot4)$ if C_α is a function of the time t only. But since all important

contributions come from the small values of $\Delta E_{\alpha\alpha'}$, the velocity of the center of wave packet will not change appreciably when an inelastic transition takes place and, therefore, if we further assume that the potential $(\alpha|V''|\alpha)$ does not depend very much on the internal state α, we can conclude that the wave packet will not deviate appreciably from that for the elastic scattering. In such a case, we can assume that C_α is a function of the time t only and does not depend on $\boldsymbol{R}$. This is the essential assumption in our semi-classical treatment. Corresponding to (8·6), we have the normalized wave packet

$$\varphi_\alpha^0(\boldsymbol{R}, t) = (2\pi)^{-3/2} \int f_\alpha^0 \exp(i\boldsymbol{k}\boldsymbol{R} - iE_k t) d\boldsymbol{k}$$

and

$$\int |\varphi_\alpha^0|^2 d\boldsymbol{R} = \int |f_\alpha^0|^2 d\boldsymbol{k} = 1.$$

Thus the equation (8·4) is transformed into

$$i\,\partial C_\alpha/\partial t = \sum_{\alpha' \neq \alpha} C_{\alpha'} \exp(i\Delta E_{\alpha\alpha'} t) \int \varphi_\alpha^{0*}(\boldsymbol{R}, t) \cdot (\alpha|V'|\alpha') \varphi_{\alpha'}^0(\boldsymbol{R}, t) d\boldsymbol{R}. \qquad (8\cdot7)$$

Strictly speaking, φ_α^0 and $\varphi_{\alpha'}^0$ do not completely coincide with each other for any finite time interval, but owing to our assumptions, these wave packets will not shift appreciably from each other during the collision time, so that we can replace them by an averaged wave packet $\varphi^0(\boldsymbol{R}, t)$. Finally, in the semi-classical treatment, $|\varphi^0(\boldsymbol{R}, t)|^2$ must be replaced by

$$\delta(\boldsymbol{R} - \boldsymbol{R}(t)).$$

$\boldsymbol{R} = \boldsymbol{R}(t)$ describes a classical orbit under the averaged potential $(\alpha|V''|\alpha)_{\mathrm{AV}}$, averaging being taken over various internal states. Generally $(\alpha|V'|\alpha)$ depends on the direction of the vector $\boldsymbol{R}$ and this dependence causes the coupling between the orbital angular momentum and the rotational angular momenta. By averaging over the various states, however, $(\alpha|V''|\alpha)_{\mathrm{AV}}$ becomes spherically symmetric which depends only on $R = |\boldsymbol{R}|$. Of course the motion of the center of wave packet will not be affected considerably even if the orbital angular momentum quantum number j changes to some extent. But if the change in j helps or reduces the inelastic transition, we can not neglect this effect and we must employ the sum

$$\sum_{j'} (\alpha, j|V'|\alpha', j') \exp\{i(\eta_{j'} - \eta_j)\}$$

instead of $(\alpha|V''|\alpha')$ for each value of j (or corresponding impact parameter b, classically). Unfortunately, it is not easy to determine the phase shifts η_j for the present case, because of the strong coupling among many states. Therefore we must employ some averaged value of this summation, e.g.

$$\left\{ \sum_{j'} (\alpha, j|V''|\alpha', j')^2 \right\}^{1/2}.$$

In order to make the resulting expression hermitian, it is convenient to modify this furthermore as follows

$$\left\{ \tfrac{1}{2} \sum_{j'} (\alpha, j|V'|\alpha', j')^2 + \tfrac{1}{2} \sum_{j'} (\alpha, j'|V'|\alpha', j)^2 \right\}^{1/2}.$$

Finally, if we consider the definite transition $\alpha_0 \rightarrow \alpha$, the initial velocity of the classical motion may be equated with the averaged value of the true initial and the true final velocities. But if we consider the many processes, some of which are exciting and some others are deactivating processes, the true initial velocity may be taken conveniently as the asymptotic velocity of the classical motion.

After these discussions, the equation (8·7) takes the following final form

$$i\, \partial C_\alpha/\partial t = \sum_{\alpha' \neq \alpha} (\alpha | V'(\boldsymbol{R}_b(t))|\alpha') C_{\alpha'} \exp(i \varDelta E_{\alpha\alpha'} t), \qquad (8\cdot8)$$

where the suffix b indicates the value of the impact parameter. We can confirm easily from this set of equations that the total probability $\sum_\alpha |C_\alpha|^2$ is independent of the time t (as is required), if the matrix $(\alpha|V'|\alpha')$ is hermitian.

For the weak coupling case, this set of equations is approximated by

$$i\, \partial C_\alpha/\partial t = (\alpha|V'|\alpha_0) C_{\alpha_0} \exp(i \varDelta E_{\alpha\alpha_0} t) \quad for \quad \alpha \neq \alpha_0 ,$$

$$C_{\alpha_0} = 1 .$$

α_0 being the initial internal state.' This gives

$$|C_\alpha(b)|^2_{t \rightarrow \infty} = |\int_{-\infty}^{\infty} (\alpha|V'(\boldsymbol{R}_b(t))|\alpha_0) \exp(i \varDelta E_{\alpha\alpha_0} t) dt|^2 \qquad (8\cdot9)$$

and the cross section for the process $\alpha_0 \rightarrow \alpha$ is given by

$$Q(\alpha_0 \rightarrow \alpha) = 2\pi \int_0^\infty |C_\alpha(b)|^2_\infty b\, db . \qquad (8\cdot10)$$

The formula (8·9), (8·10) are employed previously by Zener and others and quite recently by Widom and Bauer.[24]

If there is a group of a few states α_1, α_2, $\cdots$ which couple with the initial state α_0 strongly, and if the probabilities of the second order processes of the type : $\alpha_0 \rightarrow$ a state α_i in this group $\rightarrow$ a state outside this group are small, we may first solve the equation (8·8) for the states α_0, α_1, α_2, $\cdots$ only. This solution is performed, if necessary, by numerical calculation. Then the final probability amplitudes for the other states may be obtained by

$$|C_\alpha(b)|^2_\infty = |\int_{-\infty}^{\infty} \sum_G (\alpha|V'|G) \exp(i \varDelta E_{\alpha G} t) C_G(t) dt|^2, \quad (G \equiv \alpha_0, \alpha_1, \cdots). \quad (8\cdot11)$$

If there are many states which couple with each other considerably, we must solve a set of many coupled equations. For such a case, we shall propose here an approximate procedure of solution, which will serve to reduce the laborious calculations.

Before proceeding further, we shall write the equation (8·8) in the matrix form

$$i\, \partial \boldsymbol{C}(t)/\partial t = \boldsymbol{A}^b(t) \cdot \boldsymbol{C}(t), \qquad (8\cdot12)$$

where $\boldsymbol{C}$ is an one-column matrix and $\boldsymbol{A}^b$ is a square matrix,

$$A^b_{\alpha\alpha'} = (\alpha|V'(\boldsymbol{R}_b(t))|\alpha') \exp(i \varDelta E_{\alpha\alpha'} t),$$

all diagonal elements being lacking. Clearly this is a hermitian matrix. Now if $\boldsymbol{A}(t)$

and $\int_{-\infty}^{t} \mathbf{A}(t') dt'$ commute, we can get the formal solution of the equation $(8\cdot12)$ as follows[25]

$$\mathbf{C}(t) = \exp\left[-i\int_{-\infty}^{t} \mathbf{A}(t') dt'\right] \cdot \mathbf{C}(-\infty). \qquad (8\cdot13)$$

Here $\exp[\mathbf{M}] \equiv \sum_{0}^{\infty} \mathbf{M}^{\nu}/\nu!$, $\mathbf{M}^{\nu}$ being defined by the ordinary matrix multiplication rule. In our present case,

$$\left[\mathbf{A}(t)\int_{-\infty}^{t}\mathbf{A}(t')dt'\right]_{\alpha\alpha'} = \sum_{\alpha''}(\alpha|V'(t)|\alpha'')\exp(i\varDelta E_{\alpha\alpha''}t)$$

$$\times \int_{-\infty}^{t}(\alpha''|V'(t')|\alpha')\exp(i\varDelta E_{\alpha''\alpha'}t')dt', \qquad (8\cdot14a)$$

$$\left[\int_{-\infty}^{t}\mathbf{A}(t')dt'\cdot\mathbf{A}(t)\right]_{\alpha\alpha'} = \sum_{\alpha''}\int_{-\infty}^{t}(\alpha|V'(t')|\alpha'')\exp(i\varDelta E_{\alpha\alpha''}t')dt'$$

$$\times (\alpha''|V'(t)|\alpha')\exp(i\varDelta E_{\alpha''\alpha'}t). \qquad (8\cdot14b)$$

Generally speaking, these are not equal with each other, and, therefore, $(8\cdot13)$ can not be an exact solution of our problem. But if these two expressions are nearly equal, we may adopt $(8\cdot13)$ as an approximate solution. Now in many calculations we may put approximately that

$$(\alpha|V'(t)|\alpha') = (\alpha|v|\alpha') \cdot f(t),$$

v being independent of t. In this case, $(8\cdot14a)$ and $(8\cdot14b)$ become

$$\sum_{\alpha''}(\alpha|v|\alpha'')(\alpha''|v|\alpha')\int_{-\infty}^{t}f(t)f(t')\exp(i\varDelta E_{\alpha\alpha''}t+i\varDelta E_{\alpha''\alpha'}t')dt', \quad (8\cdot15a)$$

$$\sum_{\alpha''}(\alpha|v|\alpha'')(\alpha''|v|\alpha')\int_{-\infty}^{t}f(t)f(t')\exp(i\varDelta E_{\alpha\alpha''}t'+i\varDelta E_{\alpha''\alpha'}t)dt' \quad (8\cdot15b)$$

respectively. Here we must take the following two effects into consideration. The first is the general tendency that if the two sets of quantum numbers α and α' are extremely different from each other, the matrix element $(\alpha|v|\alpha')$ is vanishingly small. This tendency is concluded from the reasonable expectation that in the expression $(3\cdot9)$ for the intermolecular force potential, the main contribution comes from the smaller values of α, β and γ. The second effect is due to the oscillating factor $\exp(i\varDelta Et)$. If the two states α and α' being to the extremely different energy values, the integral of the form

$$\int_{-\infty}^{t} f(t')\exp(i\varDelta E_{\alpha\alpha'}t')dt'$$

must be vanishingly small, because $f(t')$ is a rather gentle function usually. Now if the first effect dominates, the main contributions to the expressions $(8\cdot15a)$, $(8\cdot15b)$ come from the region:

$$\alpha'' \simeq (\alpha+\alpha')/2$$

(this equality means, more or less symbolically, that α'' is a set of quantum numbers which is intermediate between α and α'). On the other hand, if the second effect predominates, main contribution to $(8\cdot15a)$ comes from $\alpha'' \simeq \alpha'$ and for $(8\cdot15b)$, $\alpha'' \simeq \alpha$.

In both extreme cases, the two expressions $(8\cdot15\text{a})$ and $(8\cdot15\text{b})$ are not completely coincide with each other, but, at the same time, we can not find out any essential difference between them. Thus we may conclude that $(8\cdot13)$ can not be an exact solution, but may be an approximate solution. In order to make sure, the validity of this formula must be examined for each special problem.

By employing $(8\cdot13)$, we can get the desired result

$$C_\alpha(t=+\infty) = \left[\exp\left[-i\int_{-\infty}^{\infty}\mathbf{A}(t)\,dt\right]\right]_{\alpha\alpha_0} \tag{8.16}$$

or

$$\mathbf{C}(\infty) = \mathbf{C}(-\infty) + \frac{1}{i}\int_{-\infty}^{\infty}\mathbf{A}(t')\,dt'\cdot\mathbf{C}(-\infty) - \frac{1}{2}\left\{\int_{-\infty}^{\infty}\mathbf{A}(t')\,dt'\right\}^2\cdot\mathbf{C}(-\infty) - \cdots, \tag{8.16a}$$

while the solution by the successive approximation is given by

$$\mathbf{C}(\infty) = \mathbf{C}(-\infty) + \frac{1}{i}\int_{-\infty}^{\infty}\mathbf{A}(t')\,dt'\cdot\mathbf{C}(-\infty) - \int_{-\infty}^{\infty}dt\int_{-\infty}^{t}dt'\,\mathbf{A}(t)\,\mathbf{A}(t')\cdot\mathbf{C}(-\infty) - \cdots. \tag{8.17}$$

The calculation of the series $(8\cdot16\text{a})$ is much easier to do than that of $(8\cdot17)$.

Finally, even if our approximation leads to a rather good result, we must perform such calculation for many values of impact parameter b. We can reduce, however, the number of such calculations by the same device as we employ in the modified wave number method in the next section. In another words, we can reduce the collision problem from the three-dimensional space to the head-on collison problem (of course, one-dimensional) by replacing the effective kinetic energy

$$p_0^2/2M - p_0^2 b^2/2MR^2$$

with

$$\bar{p}_0^2/2M, \quad \bar{p}_0^2 = \{1 - (b^2/R_0^2)\}p_0^2,$$

where p_0 is the initial momentum and R_0 is a suitable constant radius, the transition probabilities being neglected for $b \geq R_0$.

§ 9. Vibrational transition

As mentioned above, the distorted wave method is appropriate for calculating the cross section for the vibrational transition. Thus we may proceed as follows: In the equation $(2\cdot10)$, i.e.,

$$\{d^2/dR^2 - s(s+1)/R^2 + k_\beta^2 - 2M(\beta|V'|\beta)\}F_\beta = \sum_{\beta'\neq\beta} 2M(\beta|V'|\beta')F_{\beta'}, \tag{9.1}$$

we substitute in the right hand side the solution which is obtained by disregarding the vibrational transition. The resulting equation is for the unknown function F_β, for which the vibrational state is different from the initial state. Of course, it is desirable that the possibility of the rotational transition has already included in the zeroth order function which must be inserted in the right hand side. But to solve the problem of the rotational

transition is not easy matter as we have seen in the previous section. On the other hand, the distortion of the rotational wave functions will not considerably affect the vibrational transition, if the result is averaged over the initial rotational state. Thus it is convenient to employ the wave function corresponding to a pure rotational state, i.e., $F_{\beta'}$ are the solution of

$$\{d^2/dR^2 - j(j+1)/R^2 + k_{\beta'}{}^2 - 2M(\beta'|V'|\beta')\}F_{\beta'} = 0 \qquad (9\cdot2)$$

$$\beta' \equiv (J, \mu, n, n', j, L, l, l')$$

and are normalized asymptotically as

$$F_{\beta'} \sim s^{j l}_{J0\, m+m'} s^{ll'}_{Lnm'} \left[4\pi(2j+1)\right]^{1/2} i^j k_{\beta'}^{-1} \sin(k_{\beta'}R - j\pi/2 + \delta_{\beta'}). \qquad (9\cdot3)$$

(See § 5.)

In this way, the right hand side of $(9\cdot1)$ may be calculated and become a known function. Then the outgoing wave F_{β} can be determined as usual[18] and we have the asymptotic form of F_{β} as follows:

$$F_{\beta} \sim -2M \exp\{i(k_{\beta}R - s\pi/2 + \delta_{\beta})\}\int_0^\infty F_{\beta}^0(R) \sum_{\beta'}(\beta|V'|\beta')F_{\beta'}(R)\,dR, \qquad (9\cdot4)$$

where F_{β}^0 is the solution of the homogeneous equation which is derived from $(9\cdot1)$ by putting the right hand side zero, and is normalized asymptotically as

$$k_{\beta}^{-1} \sin(k_{\beta}R - s\pi/2 + \delta_{\beta}). \qquad (9\cdot5)$$

Thus the cross section for the transition $(n, l, m; n', l', m') \rightarrow (p, q; p', q')$ is given by

$$(k_{\beta}/k_{\beta'})\iiint\left|\sum_{JsN} F^{J\mu pp'}_{sNqq'}(R\to\infty)\, Y^{J\mu}_{sNqq'}\right|^2 d\Omega\, d\Omega_1 d\Omega_2$$

$$= (k_{\beta}/k_{\beta'})\sum_{JsN}\left|2M\int_0^\infty F_{\beta}^0(R)\sum_{\beta'}(\beta|V'|\beta')F_{\beta'}(R)\,dR\right|^2.$$

$$[\mu \equiv m+m',\ \beta = (J, \mu, p, p', s, N, q, q')]$$

Then by averaging this over m, m', we get the cross section for the inelastic process $(n, l; n', l') \rightarrow (p, q; p', q')$:

$$Q(n, l, n', l' \to p, q, p', q') = (k_{\beta}/k_{\beta'})[4M^2/(2l+1)(2l'+1)]\times$$

$$\sum_{mm'}\sum_{JsN}\left|\int_0^\infty F^{0\,J\mu pp'}_{sNqq'}\sum(Jpp's Nqq'|V'|Jnn'j Lll')F^{J\mu nn'}_{jLll'}\,dR\right|^2. \qquad (9\cdot6)$$

In principle, this formula can be applied without difficulty to any particular problem. The most difficult thing to do is the determination of the distorted waves, F_{β}^0, $F_{\beta'}$ and these functions can be obtained by solving the ordinary differential equations numerically. Nevertheless, here is a practical difficulty, because the number of differential equations to be solved is enormous. Such situation arises from the fact that a large number of the states which are defined by j, l, l' and L contribute appreciably to the collision process in the room temperature. If we know the rather exact potential V' and want to get the rigorous cross section as far as possible, we must not hesitate to find out these distorted

waves one by one.   Otherwise, however, it is desirable to seek for some simplifying pro-
cedures.

In the first place, we may neglect the difference in the potential function V' for the
different rotational states, i.e., we may employ the distorted waves which are distorted by
the averaged potential V'_0 defined in § 6.   If we adopt this simplification, the distorted
wave F_{il} does not depend on J, μ and L and depend on n, n', l, l' through the wave
number k_{il} only.   Similar situation is also obtained for F''_i.   Yet, there remain many
states with different j-values.   The best way in this stage of approximation is to carry out
the integration $\int \cdots dR$ in the above formula for some j-values and to obtain the integrals
for the other j-values by the interpolation procedure.   This interpolation will be a good
method of simplification, because it is unlikely that the integral $\int \cdots dR$ fluctuates very much
when j changes, though some fluctuations may be expected.

If we want to make further simplifications, we must give up the desire to follow the
formula (9·6).   By making such simplifications, we must also give up, more or less, to
obtain the quantitative results.   Before proceeding to discuss some such simplifications we
shall consider the general features of the vibrational transitions.

According to the kinetic theory of gases, the probability that the initial translational
energy in the thermal collision is lying in the region E, $E+dE$, is independent of the
reduced mass and the mean kinetic energy of the relative motion is $2\kappa T$ (κ is the Boltz-
mann constant; T, the absolute temperature), though the collision number depends on
the reduced mass.   Thus we may roughly consider that the initial kinetic energy is always
$2\kappa T$.   Then $k^2 = 2M \cdot E$ is proportional to the reduced mass M.   Therefore, if we compare
the equations of the form (9·2) for the two cases with different reduced masses M and
M' providing that the intermolecular force and the excitation energy are same for both
cases, we can get the similar wave functions $F(kR)$, $F(k'R)$ for the suitable pair of
quantum numbers j, j' where $j : j' = \sqrt{M} : \sqrt{M'}$.   In this case, the wave function for
the larger reduced mass oscillates more frequently than the other.   If we take the same
j-value, this contrast becomes more remarkable.   Now in contrast to this rather rapid
oscillation of the radial wave functions, the variation of the potential energy curve as the
function of R is usually gentle, at least for the collisions in the room temperature.   There-
fore, the integral in (9·6) decreases if the reduced mass increases, providing that the
radial wave functions are always normalized asymptotically to be sinusoidal functions of the
amplitude unity, or a constant amplitude independent of the reduced mass.   Of course,
the wave functions in our formula (9·6) are normalized in such a way that they are
inversely proportional to the wave number k asymptotically.   But, on the other hand,
these normalization factors are cancelled out by the factor M^2 outside the absolute square.
Thus we may conclude that the cross section increases if the reduced mass decreases providing
that the other quantities remain unchanged.   We may also say that lighter molecules are
more effective to deactivate the vibrational excitation of the particular collision partner.

We shall next consider the effect of the magnitude of ΔE, i.e., the energy change
due to the transition and also the effect of the initial wave number, providing that the
reduced mass is unchanged.   For this sake, however, we shall confine ourselves to the one-

dimensional collision problem[*], and furthermore to the intermolecular potential of the form $A \exp(-aR)$. Then the transition probability per collision is given by Jackson-Mott's formula[26]

$$P = \text{const.} \ (q_f^2 - q_i^2)^2 \sinh \pi q_f \sinh \pi q_i / (\cosh \pi q_f - \cosh \pi q_i)^2, \qquad (9 \cdot 7)$$

$$i \cdots initial, \quad f \cdots final.$$

Here, $q = 2k/a$ and $q \gg 1$ in all cases of importance. Furthermore, in the usual cases $\exp(\pi q_f) \gg \exp(\pi q_i)$ for the deactivation process and $\exp(\pi q_f) \gg \exp(\pi q_i)$ for the excitation process, because the change of the energy $\varDelta E$ is very large. Because of these situations, we have

$$P = \text{const.} \ M^2 a^{-4} (\varDelta E)^2 \exp(-2\pi |\varDelta k|/a)$$
$$|\varDelta k| = 2M |\varDelta E| / (k_i + \sqrt{k_i^2 \pm 2M |\varDelta E|}) \qquad (9 \cdot 8)$$

where the upper of the double sign corresponds to the deactivation and the lower to the excitation. Usually the exponential factor is very small and the order of magnitude of the transition probability is mainly determined by this factor. Clearly the transition probability decreases rapidly if $|\varDelta E|$ increases, while the transition probability increases if the initial wave number k_i increases. Furthermore, the transition probability increases if the quantity a increases. In many cases a is 1.5–2.5 A.U. Now, for $a=1$ and $a=3$, for example, we have extremely different transition probabilities unless $|\varDelta E|$ is very small.[3] Therefore, the transition probability is found to be very sensitive to the intermolecular force and consequently, it is very hard to get the theoretical transition probability which agrees with the experimental one within factor 2 say.

On the other hand, since a large $|\varDelta E|$ depresses the transition probability largely, it is possible that the two or even three quanta transitions may take place with larger probabilities than the one quantum transition, if $|\varDelta E|$ of the former processes are smaller than that of the latter. Thus, it is probable that in the deactivation process of the vibration, the excess energy is not transferred to the translation entirely, but some part of it is transferred to the other internal degrees of freedom. Of course, the rotational degrees of freedom can not be an acceptor of the excess energy unless the rotational level spacings are large.[**] Thus the main acceptors are the vibrational degrees of freedom, and such energy transfer between the vibrational degrees of freedom will occur appreciably if the frequencies of the two vibrations are comparable. By taking account this energy-transfer process or the coupling between the different vibrations, we can understand the fact that many of the polyatomic gases, for instance CO_2, which have many vibrational degrees of freedom, do not reveal many-relaxation-time phenomena, but behave as if these molecules have only one relaxation time.[3]

---

[*] Four atoms in the two diatomic molecules are assumed to be on a straight line.

[*] It is expected that the molecules for which the rotational level spacings are rather large, such as H_2, HCl etc., can be good acceptors of the excess energy of another molecule. In fact, according to the ultrasonic dispersion data, these molecules are found to be rather effective to deactivate the vibrational excitation of many gases.

The above formula $(9\cdot7)$, $(9\cdot8)$ can hold in the head-on collision only. In general case, in which the impact parameter is not zero we can get the general tendencies of the radial wave functions by replacing $k^2-j(j+1)/R^2$ with $\tilde{k}^2 \equiv k^2-j(j+1)/R_0^2$ in the wave equation, where R_0 is a suitable constant radius. (See the references (14) and also "the method of modified wave number" below.) Thus the effective wave number decreases rapidly with increasing j and, therefore, the transition probability decreases rapidly for the larger impact parameter.

For the vibrational transition of diatomic molecules in the room temperature, $|\varDelta E|$ is usually larger than κT. Consequently, the semi-classical treatment in which the translation is described classically is not suitable. But for many of the vibrations in the polyatomic molecules $|\varDelta E|$ becomes smaller than κT, and, at the same time, the mass of the molecule becomes large. For these larger molecules, we may apply the semi-classical method with greater and greater success.

Now we shall discuss some simplified methods to estimate the cross sections.

(i) *Utilization of the Born approximation.*

Born approximation gives usually too large cross section for the inelastic processes in the thermal collisions. This is because this method of approximation allows the approach of the two molecules without limitation. In reality, the wave function doesn't enter very much into the region inside the classical turning point, i.e., the distortion of the wave function is very important. But since the Born approximation is considerably easy to handle, it is convenient if some simple methods of correction can be found to improve the fault of Born approximation.

An interesting investigation has been done by E. Bauer.[27] According to his prescription, we must multiply a correction factor to the cross section which has been obtained by the ordinary Born approximation. This correction factor is determined by solving the corresponding one-dimensional problem for which both the plane-wave-type and the distorted-wave-type calculations of the transition matrix element can be done without difficulty. But we have no foundation that the correction factor for the three-dimensional and for the one-dimensional problems are same. In fact, the calculated cross section by Bauer was very much smaller than the experimental one. The reason for this can be understood to some extent as follows : If we decompose the wave function in the three-dimensional space into the partial waves, each of which corresponds to the definite value of the angular momentum (or impact parameter, classically), the partial wave with zero angular momentum, i.e., S-wave, will be distorted very much to the same extent as the

Fig. 4 Schematic illustration of the dependence of the cross section upon the angular momentum, incident energy being fixed.

distortion which we find in the one-dimensional problem. But for the other partial waves, the distortions are partially due to the centrifugal force, and for larger angular momenta, the distortions are mainly due to this virtual force. Such kind of distortion, however, has already been included in the Born approximation. Thus it is clear that the use of the one-dimensional correction factor in the three-dimensional problem is an overcorrection. (See the illustrating figure 4.) If we want to improve this situation, we must multiply the different correction factor for each angular momentum, or a suitable averaged correction factor to the cross section as a whole.

The other method of estimation of the cross section, utilizing the Born approximation, is the "cut-off" method. In this method, the potential is cut-off (i.e., replaced by zero) inside a suitable intermolecular distance, corresponding to the fact that the two colliding molecules can not approach without limitation. Then the calculation may be done according to the Born approximation method. As the cut-off radius, a value near the classical turning point will be adopted. But the potential curve is so steep in this region that the cross section depends rather sensitively on the cut-off radius. Therefore, the choice of the cut-off radius must be fixed empirically, from some typical examples.

(ii) *Method of the modified wave number.*

In the radial wave equations, the only difference between the S-wave case and the other cases is the appearance of the centrifugal force in the latter cases, and usually the S-wave case is much easy to treat than the others. Now if we replace $k^2 - j(j+1)/R^2$ by $k^2 - j(j+1)/R_0^2$ approximately, all equations to be solved become S-wave type, and, therefore, become easy to handle. Here R_0 is a suitable constant which may depends on k and j, in general. For larger values of j, the curve of $j(j+1)/R^2$ is rather steep and can not be replaced by a suitable constant $j(j+1)/R_0^2$ even in a narrow region near the classical turning point. But the above procedure is expected to give a rather satisfactory result, because the main contribution to the cross section comes from the smaller values of j. In fact, we have had the rather good agreement with the experimental evidences, when we have applied this method to the rotational transitions in H_2-gas.[14] Recent preliminary calculation on the vibrational transition of O_2-molecule, colliding with He-atom, has given a reasonable result too.[28]

As a concluding remark, we shall say again that for the diatomic molecules, the intermolecular force can be considered as spherical in the first approximation and then the non-spherical character may be taken into account as a perturbation, while for the polyatomic molecules, the deviation from the spherical field is essential and, therefore, the distorted wave function under the spherical potential can not be employed. For the latter case, instead, we may apply the semi-classical treatment.

§ 10. Discussions on the statistical problems. Conclusion.

In the above sections, we have treated the system consisting of two molecules only. But in almost all practical cases, we must treat an assembly of a number of molecules as a whole, and only the totality of a number of elementary processes is important. Thus some statistical procedures are required in order to connect the knowledge of the elementary

processes with the experimental data. Now in the usual theory of gases, the component molecules are assumed to be spherical particles, without internal degrees of freedom. Therefore, we must develop the theory of gases which are composed of the polyatomic molecules (including the diatomic molecules). In this paper, however, we shall confine ourselves to some brief discussions. More detailed discussions are postponed to another chance.

(1) In the pure gases, all molecules are identical except for the possibility that there may be a few kinds of isotope. Then we must apply the Bose or the Fermi statistics for these gases. When the number of the elementary particles (electrons and nucleons) in a molecule is even, the wave function of the whole system must be symmetric for an interchange of any two molecules, and for the odd number of constituent particles, the wave function must be antisymmetric. As is well known, the effect of these symmetrization is appreciable for the monatomic gases at low temperatures. For the diatomic molecules, on the other hand, we have already investigated some features of the symmetry effect in § 4. There, we could see that both even and odd values of the orbital angular momentum quantum number (j) appear in a collision process if the internal states of the two colliding molecules are different. If, however, the internal states of the two molecules are same, only the even values of j appear in the cross section formula for the Bose statistics and odd values of j for the Fermi statistics, as in the monatomic gases. We may, therefore, conclude that if the molecules in gas are distributed over many excited states, the symmetry effect will not be significant, because the probability of the encounter between the molecules with identical internal states is very small. Only one exceptional case is the hydrogen gas at low temperatures. In this case, almost all para-H_2 molecules are in the lowest rotational state while almost all ortho-H_2 molecules are in the first excited state. Of course, they are both in the lowest vibrational state. Therefore, there is a large probability of collision between the identical molecules, and thus the symmetry effect appear appreciably. We can expect a similar effect for the deuterium too. We do not enter into this problem further, but only notice that this special case has been investigated recently by K. Ohno and the present author.[17]

(2) Now the important feature of the polyatomic molecules is their non-spherical character. But, if the cross section is averaged over the initial orientations of the colliding molecules, this will not lead to an appreciable effect. When, on the other hand, the molecules are directed to some special direction to some extent by the external force, the averaged cross section becomes different according to the direction of the collision. A typical example is the *Senftleben effect*. This effect was investigated theoretically by Zernike and van Lier.[20] However, they employed a simplified geometrical consideration to estimate the collision cross section and, therefore, there remained some points which were not explained satisfactorily. The quantum mechanical method developed in the present paper will serve to improve their calculations.[*]

(3) Next, we shall consider the effects due to the inelastic collision. Two kinds

* This possibility was pointed out by Prof. Gorter in the Symposium on Molecular Physics at Nikko, Japan (September, 1953).

of effect may be distinguished. The one is the modification of the phenomena which can be found already in the monatomic gases, e.g. (probably small) corrections to the viscosity coefficient, to the diffusion constant, and other properties of gases. Since we can not pick up separately the main term and the correction term of the viscosity coefficient and of other quantities from the experimental data and since the rigorous calculation of the scattering cross section is very difficult, the corrections to these quantities are not significant, unless these quantities are affected extraordinarily by the inelastic collisions. The other kind of effects are the phenomena which can not occur in the monatomic gases, such as ultrasonic absorption due to the finite rate of the rotational or the vibrational transitions. For the effects of this kind, we can compare directly the calculated cross section of the inelastic process with the experimental data and we can also predict the results before the observations. Therefore, the theoretical investigations for this kind of effect are very useful.

In this paper, we have discussed the various features of the collisions between two diatomic molecules. In this collision problem, we are faced with two kinds of difficulties. The one is the fact that we have had no satisfactory knowledge on the intermolecular force potential, on which the collision cross section depends rather sensitively. The other is the necessity of solving the coupled set of differential equations. We have investigated mainly the latter kind of difficulties. The most easy case to handle is the weak coupling case, for which the first order approximation method of perturbation can be applied. Even in such a fortunate case, we must perform a vast volume of calculations because there are many parameters to be specified for each set of equations. Therefore, it is desirable to investigate furthermore, the approximation methods such as proposed in the last part of § 9. The validity of these proposed approximation methods will become clear when we compared the matrix elements calculated by these methods with that calculated by employing the true distorted waves.

Strong coupling case is more difficult to treat. But for such a case, the change in the translational energy during a collision is usually rather small so that we can apply the semi-classical treatment. We have proposed an approximate procedure in this direction (§ 8).

The cross sections for the elastic scatterings are mainly determined by the geometrical extension of the molecules and even rough calculations give the result of the right order of magnitude. Therefore, new calculations of these cross sections are useful, only if the reliability of the results is very high. Of course, such rigorous calculations are very difficult. Furthermore, the intermolecular potential must be determined usually from the experimental data for which only the averaged intermolecular force is responsible. Accordingly, the pure theoretical, and accurate, calculations of the elastic scattering processes may be of no significance.

Appendix I

By geometrical consideration, it is easy to see that the intermolecular potential V'' must be unchanged for the substitution operation Λ:

$$\Lambda : \quad \xi_1 = \xi_2, \quad \chi_1 \to \pi - \chi_2, \quad \chi_2 \to \pi - \chi_1 .$$

Thus

$$V'_{\alpha\beta\gamma} \equiv v'_{\alpha\beta\gamma}(\xi_1 = \xi_2, R) P_\alpha(\cos \chi_1) P_\beta(\cos \chi_2) P_\gamma(\cos \chi')$$
$$+ v'_{\beta\alpha\gamma}(\xi_1 = \xi_2, R) P_\beta(\cos \chi_1) P_\alpha(\cos \chi_2) P_\gamma(\cos \chi')$$

must be equal to

$$\Lambda V'_{\alpha\beta\gamma} \equiv v'_{\alpha\beta\gamma}(\xi_1 = \xi_2, R) P_\alpha(\cos \chi_2) P_\beta(\cos \chi_1) P_\gamma(\cos \chi')(-1)^{\alpha+\beta}$$
$$+ v'_{\beta\alpha\gamma}(\xi_1 = \xi_2, R) P_\beta(\cos \chi_2) P_\alpha(\cos \chi_1) P_\gamma(\cos \chi')(-1)^{\alpha+\beta}.$$

Other sets of (α, β, γ) don't mix with each other. By the linear independence of the Legendre polynomials, we get from the above equality

$$v'_{\alpha\beta\gamma}(\xi_1 = \xi_2, R) = (-1)^{\alpha+\beta} v'_{\beta\alpha\gamma}(\xi_1 = \xi_2, R). \qquad (3\cdot10)$$

Thus we have

$$V'_{\alpha\beta\gamma} = v'_{\alpha\beta\gamma}(\xi_1 = \xi_2, R)[P_\alpha(\cos \chi_1) P_\beta(\cos \chi_2) + (-1)^{\alpha+\beta} P_\beta(\cos \chi_1) P_\alpha(\cos \chi_2)] P_\gamma(\cos \chi')$$

and when $\alpha + \beta =$ even (odd), the expression in the bracket is symmetric (antisymmetric) function of χ_1 and χ_2 and, therefore, symmetric (antisymmetric) function of (θ_1, φ_1) and (θ_2, φ_2).

Similarly, if we employ the equality $(3\cdot12)$, we have

$$V'_{\alpha\beta\gamma} \equiv v'_{\alpha\beta\gamma}(\xi_1, \xi_2, R) P_\alpha(\cos \chi_1) P_\beta(\cos \chi_2) P_\gamma(\cos \chi')$$
$$+ v'_{\beta\alpha\gamma}(\xi_1, \xi_2, R) P_\beta(\cos \chi_1) P_\alpha(\cos \chi_2) P_\gamma(\cos \chi')$$
$$= [v'_{\alpha\beta\gamma}(\xi_1, \xi_2, R) P_\alpha(\cos \chi_1) P_\beta(\cos \chi_2)$$
$$+ (-1)^{\alpha+\beta} v'_{\alpha\beta\gamma}(\xi_2, \xi_1, R) P_\beta(\cos \chi_1) P_\alpha(\cos \chi_2)] P_\gamma(\cos \chi')$$

from which we can see that for $\alpha + \beta =$ even (odd), this expression is symmetric (antisymmetric) function of $(\xi_1, \theta_1, \varphi_1)$ and $(\xi_2, \theta_2, \varphi_2)$, and, therefore, such expression belongs to $V^+ (V^-)$.

Appendix II

The coefficients $s^{ll'}_{Lmm'}$.

When ψ_{lm} and $\psi_{l'm'}$ form the bases of the irreducible representation $D^{(l)}$, $D^{(l')}$ of the rotation group respectively, the coefficients are defined so that the linear combinations

$$\Psi_{L,m} = \sum_{m'} s^{ll'}_{L, m'\, m-m'} \psi_{lm'} \psi_{l'\, m-m'}$$

transform among themselves according to the representation $D^{(L)}$. Undetermined factor can

be chosen so that all these coefficients are real and positive. The coefficients $s_{Lmm'}^{ll'}$ are the elements of the matrix representing an unitary transformation and, therefore, have the following properties

$$\sum_\mu s_{L,\mu\,m-\mu}^{ll'}\, s_{L',\,\mu\,m-\mu}^{ll'} = \delta_{LL'} ,\tag{A2·1}$$

$$\sum_L s_{L,\mu\,m-\mu}^{ll'}\, s_{L,\,\mu'\,m-\mu'}^{ll'} = \delta_{\mu\mu'} .\tag{A2·2}$$

Sometimes the notation $(ll'\,m\,m'|ll'\,L\,m+m')$ is used instead of $s_{L,mm'}^{ll'}$.

The cross section for viscosity.

The cross section which is effective for the viscosity, is given by the formula

$$Q_{vis} = \sum_{jj'JJ''L,l'} (E_{JjL}^{lml'm'}{}^{*}\, E_{j''j'l'L'}^{lml'm'})_{R\to\infty} \iint \mathbf{Y}_{jl,ll'}^{J\,m+m'}{}^{*}\, \mathbf{Y}_{j'l',lll'}^{J'\,m+m'}\sin^2\theta\, d\Omega\, d\Omega_1 d\Omega_2 .\tag{A2·3}$$

(As to the notations, see $(5\cdot2)$.) Since, by definition,

$$\mathbf{Y}_{jl,ll'}^{J\mu} \equiv \sum_\nu s_{J,\mu-\nu\,\nu}^{jL}\, Y_{j\,\mu-\nu}(\theta\Phi) \sum_{\nu'} s_{L,\nu'\,\nu-\nu'}^{ll'}\, Y_{l\nu'}(\theta_1\varphi_1)\, Y_{l'\,\nu-\nu'}(\theta_2\varphi_2),$$

we can easily integrate over $d\Omega_1 d\Omega_2$ and have

$$\iiint \cdots = \delta_{LL'} \sum_\nu s_{J,\mu-\nu\,\nu}^{jL}\, s_{J'',\mu-\nu\,\nu}^{j'L} \int Y_{j\,\mu-\nu}{}^{*}(\theta\Phi)\, Y_{j'\,\mu-\nu}(\theta\Phi)\sin^2\theta\, d\Omega$$

$$= \delta_{LL'} \sum_\nu s_{Jv\,\mu-\nu}^{jL}\, s_{J''\,v\,\mu-\nu}^{j'L} \int Y_{j\nu}{}^{*}(\theta\Phi)\, Y_{j'\nu}(\theta\Phi)\sin^2\theta\, d\Omega\tag{A2·4}$$

where we have used the orthonormality of the spherical harmonics and $(A2\cdot1)$. Next, by making use of the well-known formula

$$\iint Y_{j\nu}{}^{*}(\theta,\varphi)\, Y_{j'\nu}(\theta,\varphi)\sin^3\theta\, d\theta\, d\varphi = \frac{2(j^2+j-1+\nu^2)}{(2j-1)(2j+3)}\cdot\delta_{jj'}$$

$$-\frac{1}{2j+3}\sqrt{\frac{(j+1+\nu)(j+1-\nu)(j+2+\nu)(j+2-\nu)}{(2j+1)(2j+5)}}\,\delta_{j+2,\,j'}$$

$$-\frac{1}{2j-1}\sqrt{\frac{(j-1+\nu)(j-1-\nu)(j+\nu)(j-\nu)}{(2j+1)(2j-3)}}\,\delta_{j-2,j'} .\tag{A2·5}$$

and by remembering the asymptotic form of $E_{JjL}^{lml'm'}$ (see $(5\cdot2)$.) the cross section Q_{vis} becomes : $(\mu\equiv m+m')$

$$Q_{vis} = \frac{\pi}{k^2} \sum_{j,J,J''L,\nu} s_{Jv\,\mu-\nu}^{jL}\, s_{J,0\mu}^{jL} (s_{L,mm'}^{ll'})^2 \left[s_{J''\,v\,\mu-\nu}^{jL}\, s_{J''0\mu}^{jL}\, \frac{2(2j+1)(j^2+j-1+\nu^2)}{(2j-1)(2j+3)}|S_{JjL}^{lml'm'}-1|^2 \right.$$

$$-s_{J''\,v\,\mu-\nu}^{j+2,\,L}\, s_{J''0\mu}^{j+2,\,L}\, \frac{\sqrt{(j+1+\nu)(j+1-\nu)(j+2+\nu)(j+2-\nu)}}{2j+3}$$

$$\times (S_{JjL}^{lml'm'}{}^{*}-1)(S_{J''\,j+2,\,L}^{lml'm'}-1)$$

$$\left. -s_{J''\,v\,\mu-\nu}^{j-2,\,L}\, s_{J''0\mu}^{j-2,\,L}\, \frac{\sqrt{(j-1+\nu)(j-1-\nu)(j+\nu)(j-\nu)}}{2j-1}\, (S_{JjL}^{lml'm'}{}^{*}-1)(S_{J''\,j-2,\,L}^{lml'm'}-1) \right].$$

When we sum up over J and J', the second and the third terms become complex conjugate with each other.   Therefore, we have

$$Q_{vts} = \frac{2\pi}{k^2} \sum_{j,l,l',l,\nu} \frac{1}{2j+3} S_{Jl\nu}^{jL}{}_{\mu-\nu} S_{Jl0\mu}^{jL} (S_{l,mm'}^{jl'})^2 \left[S_{,l'}^{jL}{}_{\nu}{}_{\mu-\nu} S_{,l'0\mu}^{jL} \right.$$

$$\times \frac{(2j+1)(j^2+j-1+\nu^2)}{2j-1} |S_{Jjl}^{lml'm'} - 1|^2$$

$$- S_{,l'}^{j+2,L}{}_{\nu}{}_{\mu-\nu} S_{,l'0\mu}^{j+2,L} \sqrt{(j+1+\nu)(j+1-\nu)(j+2+\nu)(j+2-\nu)}$$

$$\left. \times \mathrm{Re} \left\{ (S_{Jjl}^{lml'm'*} - 1)(S_{Jj+2,L}^{lml'm'} - 1) \right\} \right]. \qquad (A2 \cdot 6)$$

Appendix III

Here, we shall prove the two relations which have been used in § 6.

(i)   *Independence of the matrix elements of the intermolecular potential V' upon the quantum number μ* which defines the z-component of the total angular momentum of the whole system.

Since our system consists of four atoms, the configuration can be expressed completely by 12 coordinates, which may be classified as follows :

I :   three coordinates which define the position of the center-of-gravity of the system,

II :   three coordinates R, ξ_1, ξ_2,   (See Fig. 2.)

III :   three coordinates χ_1, χ_2, $\omega_1 - \omega_2$, which define the relative orientation of the two molecules,

IV :   three coordinates which define the orientation of the system as a whole.

Now we are considering the matrix elements $(J\mu jLll' | V' | J'\mu'j' L'\lambda\lambda')$ in the subspace (III) + (IV) and wave function can be written in this subspace as [4][6]

$$\psi_{J\mu}^{jLll'} = \sum_s D^J(\mathcal{R})_{s\mu} \chi_{Js}^{jLll'}(\chi_1, \chi_2, \omega_1 - \omega_2)$$

where $\mathcal{R}$ is the rotation of coordinate system which bring the system into a standard configuration.[6]   Thus we have

$$(J\mu jLll' | V' | J'\mu'j' L'\lambda\lambda')$$

$$= \sum_{ss'} \int dv_{III} \int dv_{IV} D^{J*}(\mathcal{R})_{s\mu} D^{J'}(\mathcal{R})_{s'\mu'} \chi_{Js}^{jLll'*} \chi_{J's'}^{j'L'\lambda\lambda'} V'$$

$$= \sum_{ss'} \int dv_{III} \chi_{Js}^{jLll'*} V' \chi_{J's'}^{j'L'\lambda\lambda'} \int dv_{IV} D^{J*}(\mathcal{R})_{s\mu} D^{J'}(\mathcal{R})_{s'\mu'}.$$

The first integral is independent of $\mathcal{R}$ while the second integral becomes $(1/2J+1) \times \delta_{JJ'} \delta_{ss'} \delta_{\mu\mu'} \int dv_{IV}$ by the property of the representation of rotation group.   Thus we obtain that the above matrix element is

$$(2J+1)^{-1} \delta_{JJ'} \delta_{\mu\mu'} \sum_s \int dv_{III} \chi_{Js}^{jLll'*} V' \chi_{Js}^{j'L'\lambda\lambda'} \int dv_{IV}$$

the value of which is independent of μ, so long as $\mu = \mu'$.

(ii)　*Proof of* $\displaystyle\sum_{mm'}\sum_{JL}\big(s^{jL}_{J0\,m+m'}\,s^{ll'}_{Lmm'}\big)^{2}\,\varDelta V^{JL}_{jll'}=0$ $\qquad\cdots(6\cdot5)$

From the definition, we get

$$V_0=(1/2l+1)(1/2l'+1)\sum_{mm'}(j0lml'm'|V'|j0lml'm')$$

$$=(1/2l+1)(1/2l'+1)\sum_{mm'}\sum_{JJ'LL'}s^{jL}_{J0\,m+m'}\,s^{jL'}_{J'0\,m+m'}\,s^{ll'}_{Lmm'}\,s^{ll'}_{L'mm'}\,(J\mu jLll'|V'|J'\mu jL'll').$$

Since the elements in the summation do not depend on $\mu=m+m'$ and vanish for $J\neq J'$ as just proved in (i), the summation on m and m' can be carried out as

$$\sum_{mm'}s^{jL}_{J0\,m+m'}\,s^{jL'}_{J0\,m+m'}\,s^{ll'}_{Lmm'}\,s^{ll'}_{L'mm'}$$

$$=\sum_{\mu}s^{jL}_{J0\mu}\,s^{jL'}_{J0\mu}\sum_{\nu}s^{ll'}_{L\nu\,\mu-\nu}\,s^{ll'}_{L'\,\nu\,\mu-\nu}$$

$$=\sum_{\mu}(s^{jL}_{J0\mu})^{2}\cdot\hat{\delta}_{LL'}\,,$$

where we have used (A2·1). Thus we have

$$V_0=(1/2l+1)(1/2l'+1)\sum_{JL\mu}(s^{jL}_{J0\mu})^{2}(JjLll'|V'|JjLll').\qquad(\mathrm{A}3\cdot1)$$

Now, from the definition,

$$\varDelta V^{JL}_{jll'}=(JjLll'|V'|JjLll')-V_0$$

and, therefore, we get

$$\sum_{mm'}\sum_{JL}\big(s^{jL}_{J0\,m+m'}\,s^{ll'}_{Lmm'}\big)^{2}\varDelta V^{JL}_{jll'}$$

$$=\sum_{JL}\Big\{\sum_{\mu\nu}\big(s^{jL}_{J0\mu}\,s^{ll'}_{L\nu\,\mu-\nu}\big)^{2}\Big\}(JjLll'|\,l''\,|JjLll')-V_0\sum_{mm'}\sum_{JL}\big(s^{jL}_{J0\,m+m'}\,s^{ll'}_{Lmm'}\big)^{2}$$

$$=\sum_{JL\mu}(s^{jL}_{J0\mu})^{2}(JjLll'|V'|JjLll')-V_0(2l+1)(2l'+1)\qquad\text{by }(\mathrm{A}2\cdot1),\ (\mathrm{A}2\cdot2)$$

$$=0\quad\text{by }(\mathrm{A}3\cdot1).$$

References

1) C. Zener, Phys. Rev. **37** (1931) 556; Proc. Cambr. Phil. Soc. **29** (1932), 136.
2) R. N. Schwartz, Z. I. Slawsky and K. F. Herzfeld, J. Chem. Phys. **20** (1952), 1591.
3) K. Takayanagi, Prog. Theor. Phys. **8** (1952), 111.
4) C. F. Curtiss, J. O. Hirschfelder and F. T. Adler, J. Chem. Phys. **18** (1950), 1638.
5) C. F. Curtiss and F. T. Adler, J. Chem. Phys. **20** (1952), 249.
6) C. F. Curtiss, J. Chem. Phys. **21** (1953), 1199, 2045.
7) E. Wigner, *Gruppentheorie und ihre Anwendung auf die Quantenmechanik der Atomspektren.*
8) H. Margenau, Phys. Rev. **63** (1943), 131.
9) J. de Boer, Physica **9** (1942), 363.
10) A. A. Evett and H. Margenau, Phys. Rev. **90** (1953), 1021.
11) A. S. Roy and M. E. Rose, Proc. Roy. Soc. **A149** (1935), 511.
12) E. U. Condon and G. H. Shortley, *Theory of Atomic Spectra* (1935).
13) G. H. Shortley and B. Fried, Phys. Rev. **54** (1938), 739.
14) K. Takayanagi, Prog. Theor. Phys. **8** (1952), 497; K. Takayanagi and T. Kishimoto, Prog. Theor. Phys. **9** (1953), 578; **10** (1953), 369 (errata),

15) R. H. Fowler and E. A. Guggenheim, *Statistical Thermodynamics* (1939), p. 718.

16) H. S. W. Massey and E. H. S. Burhop, *Electronic and Ionic Impact Phenomena* (1952), 460.

17) Progress Report No. 2 (July, 1953) of " Research Group for the study of Atomic and Molecular Structure " p. 25. A little more improved discussions were reported at Nikko Symposium on Molecular Physics on 12th, September, 1953 and the content of the lecture will be published soon in the Proceedings of Symposium.

18) N. F. Mott and H. S. W. Massey, *Theory of Atomic Collision*, (1949), Chapter VI.

19) E. C. G. Stueckelberg, Helv. Phys. Acta **5** (1932), 369.

20) H. S. W. Massey and C. B. O. Mohr, Proc. Phys. Soc. **A65** (1952), 845.

21) N. F. Mott and H. S. W. Massey, *Theory of Atomic Collisions* (1949), p. 153.

22) E. S. Stewart, Phys. Rev. **69** (1946), 632.

23) J. E. Rhodes Jr. Phys. Rev. **70** (1946), 932.

24) J. E. Frame, Proc. Cambr. Phil. Soc. **27** (1931), 511 ; **33** (1937), 115.
     N. F. Mott, Proc. Cambr. Phil. Soc. **27** (1931), 553.
     C. Zener, Phys. Rev. **38** (1931), 277 ; Proc. Cambr. Phil. Soc. **29** (1932), 136.
     E. J. Williams, Proc. Roy. Soc. **A139** (1933), 163.
     Benjamin Widom and S. H. Bauer, J. Chem. Phys. **21** (1953), 1670.

24) N. Arley, *On the Theory of Stochastic Processes and their Application to the Theory of Cosmic Radiation* (1943), Chapter 2.

26) J. M. Jackson and N. F. Mott, Proc. Roy. Soc. **A137** (1932), 703. See also A. F. Devonshire, Proc. Roy. Soc. **A158** (1937), 269 in which the Morse type intermolecular potential was adopted.

27) E. Bauer, Phys. Rev. **84** (1951), 315 ; **85** (1952), 277.

28) K. Takayanagi and S. Kaneko, Progress Report No. 3 (1954) of " Research Group for the Study of Atomic and Molecular Structure,"

29) F. Zernik and C. Van Lier, Physica **6** (1639), 961.

Note added in proof : Detailed report on our work in the reference (28) will be published soon in the Science Reports of the Saitama University.

Two Simple Stochastic Models of Cascade Multiplication

Alladi RAMAKRISHNAN and S. K. SRINIVASAN

University of Madras, S. India

Numerical results are obtained in regard to two simple stochastic models of cascade multiplication. It is shown that these models reveal many interesting features characteristic of more complicated multiplicative processes like the well known stochastic process of cosmic ray cascades.

§ 1.   Introduction

In a previous contribution to this Journal, Ramakrishnan and Mathews[1] presented their latest numerical calculations on the well known stochastic problem of cosmic radiation. It belongs to the general class of multiplicative processes where one has to deal with a stochastic variable representing the number of particles distributed in a continuous infinity of states characterised by a parameter say E (e.g. E may denote energy). In any mathematical treatment using the exact transition probabilities, the numerical evaluation of the results is very difficult and the essential physical features of the process are sometimes obscured by the complexity of the problem. Therefore it seemed worthwhile to propose and solve stochastic problems relating to two simple multiplicative models of cascade processes yielding elegant analytical results amenable to numerical computation. *It is shown from these numerical results that there is an essential difference between processes involving finite and infinite total differential cross-sections.* Our results seem to confirm indirectly the extensive calculations of Ramakrishnan and Mathews.

The fluctuation problem of cosmic radiation has attracted the attention of both mathematicians and physicists and many papers have been published on such stochastic processes of " continuous parametric systems." But in not one of them has this simple but essential difference between processes involving finite and infinite total cross-sections been noticed or stressed.

§ 2.   Statement of the problem

We are given that the probability per unit thickness of matter that a particle of energy E splits up into two, one of which has an energy between Eq and $E(q+dq)$ and the other between $E(1-q)$* and $E(1-q-dq)$ is $2\omega(q)dq, (q<1)$. The factor two occurs because we treat the particles as indistinguishable. Our object is to calculate the mean and mean square number of particles above an energy E_C at thickness t if a particle of initial energy E_0 is incident on matter.

---

* We here assume that sum of the energies of the ' offsprings ' is equal to the energy of the ' parent.' The case when this is not so is discussed by (R) in an earlier paper, Proc. Camb. Phil. Soc. **48** (1952), 451.

§ 3. The solution of the problem

Defining $f_1(E\,;t)$ and $f_2(E_1, E_2\,;t)$ as the product densities of degree one and two of particles and using a method now familiar, we obtain

$$\frac{\partial f_1(E\,;t)}{\partial t} = -f_1(E\,;t)\int_0^1 \omega(q)\,dq + 2\int_{E/E_0}^1 f_1\left(\frac{E}{q}\,;t\right)\frac{\omega(q)}{q}\,dq, \tag{1}$$

$$\frac{\partial f_2(E_1, E_2\,;t)}{\partial t} = -2f_2(E_1, E_2\,;t)\int_0^1 \omega(q)\,dq + \int_{E_1/E_0}^1 2f_2\left(\frac{E_1}{q}, E_2\,;t\right)\frac{\omega(q)}{q}\,dq$$

$$+\int_{E_2/E_0}^1 2f_2\left(E_1, \frac{E_2}{q}\,;t\right)\frac{\omega(q)}{q}\,dq + 2f_1(E_1+E_2\,;t)\frac{\omega(q')}{q'}, \tag{2}$$

where

$$q' = E_1/(E_1+E_2) \tag{3}$$

$\lambda = \int_0^1 \omega(q)\,dq$, the total probability per unit thickness that a particle splits up into two, is assumed to be finite. The case when λ is infinite will be discussed presently.

Defining the Mellin's transformations,

$$C_1(s\,;t) = \int_0^\infty f_1(E\,;t)E^{s-1}\,dE, \tag{4}$$

$$C_2(r, s\,;t) = \int_0^\infty \int_0^\infty f_2(E_1, E_2\,;t)E_1^{r-1}E_2^{s-1}\,dE_1\,dE_2, \tag{5}$$

equations (1) and (2) reduce to the differential equations,

$$\partial C_1(s\,;t)/\partial t = -C_1(s\,;t)\{\lambda - a(s)\}, \tag{6}$$

$$\partial C_2(r, s\,;t)/\partial t = -C_2(r, s\,;t)\{2\lambda - a(r) - a(s)\} + \alpha(r, s)C_1(r+s-1\,;t) \tag{7}$$

where

$$a(s) = 2\int_0^1 \omega(q)q^{s-1}\,dq, \quad a(1) = 2\lambda, \quad \alpha(r, s) = 2\int_0^1 \omega(q)q^{r-1}(1-q)^{s-1}\,dq. \tag{8}$$

Since at $t=0$ there is one particle with energy E_0,

$$f_1(E\,;0) = \delta(E_0 - E), \quad f_2(E_1, E_2\,;0) = 0, \tag{9}$$

where δ is the Dirac delta function. Consequently

$$C_1(s\,;0) = E_0^{s-1}, \quad C_2(r, s\,;0) = 0. \tag{10}$$

Solving (6) and (7) we obtain

$$C_1(s\,;t) = E_0^{s-1}e^{-A_s t} \tag{11}$$

$$C_2(r, s\,;t) = E_0^{r+s-2}\frac{\alpha(r, s)\left[e^{-A_{r+s-1}t} - e^{-(A_r+A_s)t}\right]}{A_r + A_s - A_{r+s-1}} \tag{12}$$

where

$$A_s = \lambda - a(s) \tag{13}$$

f_1 and f_2 are obtained by using the inversion formulae

$$f_1(E\,;t) = \frac{1}{2\pi i}\int_{\sigma-i\infty}^{\sigma+i\infty} C_1(s\,;t)E^{-s}\,ds \tag{14}$$

$$f_2(E_1, E_2; t) = \frac{1}{(2\pi i)^2} \int_{\sigma-i\infty}^{\sigma+i\infty} \int_{\sigma-i\infty}^{\sigma+i\infty} C_2(r, s; t) E_1^{-r} E_2^{-s} dr ds. \quad (15)$$

According to the theory of product densities the mean number of particles above the energy E_C is given by

$$\mathcal{E}\{N(E_C; t)\} = \int_{E_C}^{E_0} f_1(E; t) dE = \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} e^{\phi(s; t)} ds, \quad (16)$$

where

$$\phi(s; t) = y(s-1) - A_s t - \log(s-1), \quad y = \log(E_0/E_C). \quad (17)$$

The mean square number of particles above the energy E_C is given by

$$\mathcal{E}\{N^2(E_C; t)\} = \mathcal{E}\{N(E_C; t)\} + \int_{E_C}^{E_0} \int_{E_C}^{E_0} f_2(E_1, E_2; t) dE_1 dE_2$$

$$= \mathcal{E}\{N(E_C; t)\} + \frac{1}{(2\pi i)^2} \int_{\sigma-i\infty}^{\sigma+i\infty} \int_{\sigma-i\infty}^{\sigma+i\infty} e^{\psi(r, s; t)} dr ds, \quad (18)$$

where

$$\psi(r, s; t) = y(r+s-2) + \log(\alpha(r, s))$$

$$+ \log\left[\frac{e^{-A_{r+s-1}t} - e^{-(A_r+A_s)t}}{A_r + A_s - A_{r+s-1}}\right] - \log[(s-1)(r-1)]. \quad (19)$$

Noting that $\mathcal{E}\{N(E_C; t)\}$ and $\mathcal{E}\{N^2(E_C; t)\}$ are functions only of E_0/E_C or $y = \log(E_0/E_C)$, we can write them as $\mathcal{E}\{N(y; t)\}$ and $\mathcal{E}\{N^2(y; t)\}$. They are evaluated for various values of y and t using the well known saddle point method according to which

$$\mathcal{E}\{N(y; t)\} = \frac{e^{\phi(s_0; t)}}{\left[2\pi \frac{\partial^2 \phi(s; t)}{\partial s^2}\right]_{s=s_0}^{1/2}} \quad (20)$$

where s_0 is the saddle point satisfying the equation

$$[\partial\phi(s; t)\partial s]_{s=s_0} = 0. \quad (21)$$

The saddle point formula corresponding to (19) is

$$\frac{1}{(2\pi i)^2} \int_{\sigma-i\infty}^{\sigma+i\infty} \int_{\sigma-i\infty}^{\sigma+i\infty} e^{\psi(r, s; t)} dr ds = \left\{\frac{e^{\psi(r, s; t)}}{2\pi\left[\left(\frac{\partial^2 \psi}{\partial r^2}\right)^2 - \left(\frac{\partial^2 \psi}{\partial r \partial s}\right)^2\right]^{1/2}}\right\}_{r=s=s_0}, \quad (22)$$

where s_0 is the saddle point satisfying the equation

$$(\partial\psi/\partial s)_{r=s=s_0} = 0, \quad (\partial\psi/\partial r)_{r=s=s_0} = 0. \quad (23)$$

We realise that the complexity of the problem depends essentially on the " differential cross-section " $\omega(q)$.

Model A (Finite cross-section)

The model is defined by

$$\omega(q)\,dq \equiv \lambda\,dq \tag{24}$$

i.e.
$$\int_0^1 \omega(q)\,dq = \lambda \tag{25}$$

i.e. the probability that a particle of energy E splits up into two, one of which has an energy E' is $2dE'/E$. We shall now show that the simple model reveals many interesting features characteristic of the general class of stochastic processes mentioned in the introduction.

1. $\mathcal{E}\{N(y;t)\}$ has a maximum for a given y at a certain value of $t = t_y$ such that the saddle point corresponding to t_y and y is $s_0 = 2$.

$$t_y = 2(y-1)/\lambda, \tag{26}$$

$$\mathcal{E}\{N(y;t_y)\} = e^y/(2\pi y)^{1/2} . \tag{27}$$

2. In the present case where $\omega(q)\,dq \equiv \lambda\,dq$ it is possible to obtain an explicit expression for $\mathcal{E}\{N(y;t)\}$ by evaluating the complex integral (14) using the theory of residues.

$$\mathcal{E}\{N(y;t)\} = e^{-\lambda t} + \sum_{n=1}^{\infty} e^{-\lambda t} \frac{(2\lambda t)^n}{n!} \int_0^y \frac{e^{-y'} y'^{n-1}}{(n-1)!}\,dy' . \tag{28}$$

We, however, found that for the numerical computation of $\mathcal{E}\{N(y;t)\}$ for finite y the saddle point formula (20) was more suitable than (28).

3. The complex integrals (14) and (15) cannot be evaluated easily for $y \to \infty$ i.e. when we wish to calculate the mean and mean square of the total number of particles. This is easier done by integrating the fundamental equations (1) and (2) over the entire energy range when we obtain

$$\mathcal{E}\{N(y;t)\} \to e^{\lambda t}, \quad \mathcal{E}\{N^2(y;t)\} \to e^{2\lambda t} - e^{\lambda t} \quad \text{as} \quad y \to \infty, \tag{29}$$

in accordance with the well known Furry law. This is true for any general $\omega(q)$ provided λ is finite. In the case when $y \to \infty$ it is possible to compute $\pi(N;t)$ the probability that there are N particles at t. π satisfies the equation

$$\partial \pi(N;t)/\partial t = \lambda [(N-1)\pi(N-1;t) - N\pi(N;t)] \tag{30}$$

whence

$$\pi(N;t) = \frac{1}{\mathcal{E}\{N\}} \left[1 - \frac{1}{\mathcal{E}\{N\}} \right]^{N-1}. \tag{31}$$

We note that if the stochastic variable N has a Furry distribution

$$\sigma^2 = \mathcal{E}\{N^2\} - (\mathcal{E}\{N\})^2 = (\mathcal{E}\{N\})^2 - \mathcal{E}\{N\}. \tag{32}$$

For a Poisson distribution $\sigma^2 = \mathcal{E}\{N\}$. It has become customary in the case of multiplicative stochastic processes of the type we are discussing to compare $\sigma^2(y;t)$ i.e. $\mathcal{E}\{N^2(y;t)\} - (\mathcal{E}\{N(y;t)\})^2$ with $\mathcal{E}\{N(y;t)\}$ and $(\mathcal{E}\{N(y;t)\})^2 - \mathcal{E}\{N(y;t)\}$ which will be denoted by σ_P^2 and σ_F^2 respectively for the obvious reason that they are the values we should expect for the mean square deviation if $N(y;t)$ were to have a Poisson or Furry distribution. Tables of $\mathcal{E}\{N(y;t)\}, \mathcal{E}\{N^2(y;t)\}, \sigma^2(y;t), \sigma^2/\sigma_P^2, \sigma^2/\sigma_F^2$ for various

values of y and t are given at the end of the paper. λ is chosen to be unity without loss of generality.

Comparing with the Poisson distribution, σ^2/σ_P^2 increases with y rapidly. Comparison with the Furry distribution is meaningful only when $\mathcal{E}\{N(y;t)\}$ is, say, above 5. Beyond a certain value of y, $\sigma^2/\sigma_F^2 < 1$, but increases appreciably with y. As $y \rightarrow \infty$, $\sigma^2/\sigma_F^2 \rightarrow 1$.

Model B.

We now choose the following cross section

$$\omega(q) = \frac{1}{4}\left(\frac{1}{q} + \frac{1}{1-q}\right). \tag{33}$$

We immediately note that ω is symmetrical in q and $1-q$ and the total cross-section $\int_0^1 \omega(q)\,dq$ is infinite. Such a difficulty arose in the stochastic problem of cosmic ray cascades since the total Bethe-Heitler cross-section for the radiation of a photon by an electron is infinite. Fortunately, for the calculation of the moments of the number of particles above a certain energy i.e. for finite values of y, an infinite cross-section does not present any difficulty since in equation (1) the two terms on the right hand side together converge in the Cauchy sense while each of them is infinite. Or in other words,

$$\left[\int_0^1 \omega(q)\,dq - 2\int_0^1 \omega(q)q^{s-1}\,dq\right] \tag{34}$$

is convergent though each of the terms inside the square brackets diverges. The above expression is exactly equal to

$$\frac{1}{2}\left[\frac{d}{ds}\log\overline{|s-1|} + \gamma\right], \tag{35}$$

where γ is the Euler-Mascheroni constant.

$$\alpha(r,s) = 1/2 \cdot [B(r-1,s) + B(r,s-1)] = 1/2 \cdot B(r-1,s-1) \tag{36}$$

where B is the Beta function.

For a given y, $\mathcal{E}\{N(y;t)\}$ has a maximum value at t_y and the saddle point s_y corresponding to y and t_y is given by

$$\left[\frac{d}{ds}\log\overline{|s-1|} + \gamma\right]_{s_y} = 0. \tag{37}$$

In the case of an infinite cross-section, $\mathcal{E}\{N(y;t)\}$ tends to infinity as $y \rightarrow \infty$. Thus the Furry distribution is approached only as the mean tends to infinity. It is to be expected therefore that for a given t, for finite values of y, σ^2/σ_F^2 is less than unity beyond a certain value of y and increases in a much slower manner than in the case of finite total cross-section. (Comparison with σ_F^2 is meaningful only when the mean number of particles is, say, above five). Nevertheless, σ^2/σ_P^2 increases with y for a given t. It is many times the Poisson value even at the cascade maximum.

. It is noteworthy that this simple model reveals all the interesting features of the electron cascade. Though in the case of the cosmic ray cascade there are two types of particles, electrons and photons, it was shown by one of us[2] that provided we assume the total cross-section for radiation is finite, the total number of electrons has an approximately Furry distribution (for all but small t) as in the case of a cascade consisting of a single type of particle.

When the stochastic features of two multiplicative processes are to be compared, we must choose the units of thickness in each case in such a manner that for a given y we obtain for the two cases, mean numbers of the same order for the same numerical value of t. The change in the unit of thickness means a corresponding change in the differential cross-section per unit thickness. For example, if we are asked to compare a process with cross-section $\dfrac{1}{q} + \dfrac{1}{1-q}$ with that of Model A, we choose the unit of thickness of the former to be 1/4 that of the Model A, so that its " new " cross-section is $\dfrac{1}{4}\left(\dfrac{1}{q} + \dfrac{1}{1-q}\right)$.

Finally, we wish to enunciate a lemma which we believe to be true but for which we have not been able to provide a mathematical proof.

Lemma :   In the case of any multiplicative process defined by equations (1) and (2) for a given t, for all finite y greater than a particular value (which depends on t, and increases with y) $(\sigma^2(y;t)/\sigma_F^2) < 1$, but increases with y, and tends to unity as $y \to \infty$.

<div align="center">

Model A (cross section : finite)

Table I.

$\varepsilon\{N(y;t)\}$

$[\varepsilon\{N(y;t)\}]^2$ is given in brackets.

</div>

y \ t	1	2	3	4	5	6	8
2	1.74 (3.02)	2.08 (4.34)	1.98 (3.90)	1.94 (3.75)			
3	2.11 (4.46)	3.24 (10.5)	4.14 (17.2)	4.63 (21.4)	4.67 (21.9)	4.27 (18.2)	3.21 (10.3)
4		4.29 (18.4)	6.51 (42.4)	8.56 (73.3)	10.1 (101)	10.9 (119)	10.4 (108)
5		5.13 (26.3)	8.94 (79.9)	13.5 (181)	18.1 (327)	22.7 (515)	26.5 (701)
6			11.3 (127)	18.9 (358)	28.4 (808)	38.7 (1500)	56.8 (3230)
7				24.5 (600)	40.5 (1640)	60.7 (3690)	107 (11500)
8					53.7 (2880)	87.6 (7670)	182 (33100)
9						120 (14300)	283 (79800)
10							415 (173×10^5)
11							573 (328×10^5)

Table II.

$$\varepsilon\{N^2(y;t)\}$$

$\sigma^2(y;t)$ is given in brackets.

t y	1	2	3	4	5	6	8
2	3.92 (.90)	4.38 (.04)	7.78 (3.88)	6.63 (2.88)			
3	5.83 (1.38)	13.3 (2.8)	20.2 (3.1)	24.3 (2.9)	24.3 (2.5)	21.2 (3.0)	13.6 (3.3)
4		25.4 (7.0)	54.1 (11.7)	87.3 (14.0)	115 (14)	129 (11)	116 (7.2)
5		40.4 (14.1)	110 (30)	232 (51)	390 (63)·	577 (62)	741 (40)
6			184 (57)	486 (130)	1080 (280)	1770 (280)	3550 (320)
7				860 (260)	2210 (570)	4630 (940)	12900 (1500)
8					4040 (1160)	102×10^2 (25×10^2)	398×10^2 (67×10^3)
9						194×10^2 (51×10^2)	101×10^3 (21×10^3)
10							225×10^3 (52×10^3)
11							446×10^3 (118×10^3)

Table III.

$$\sigma^2/\sigma_F^2$$

σ^2/σ_P^2 is given in brackets.

t y	1	2	3	4	5	6	8
2	.71 (.52)	.028 (.017)	2.0 (2.0)	1.6 (1.5)			.37 (.88)
3	.59 (.65)	.44 (.85)	.23 (.74)	.17 (.63)	.14 (.51)	.22 (.71)	.073 (.69)
4		.47 (1.6)	.33 (1.8)	.22 (1.6)	.16 (1.5)	.10 (.99)	.059 (1.5)
5		.67 (2.8)	.42 (3.3)	.30 (3.7)	.20 (2.9)	.13 (2.7)	.10 (5.7)
6			.49 (5.1)	.38 (6.8)	.28 (7.7)	.19 (7.1)	.13 (14)
7				.45 (11)	.35 (14)	.26 (16)	.20 (37)
8					.41 (22)	.33 (28)	.27 (75)
9						.36 (42)	.30 (130)
10							.36 (200)
11							

A. Ramakrishnan and S. K. Srinivasan

Model B (cross section: Infinite)
Table IV.
$$\varepsilon\{N(y;t)\}$$
$[\varepsilon\{N(y;t)\}]^2$ is given in brackets.

t / y	2	3	4	5	6	7	8
2	1.96 (3.83)	1.66 (2.76)	1.26 (1.59)				
3	3.79 (14.3)	4.45 (19.8)	3.93 (15.5)	3.26 (10.6)	2.45 (6.00)	1.70 (2.89)	
4	6.51 (42.4)	8.63 (74.5)	9.59 (92.0)	9.38 (88.0)	8.35 (69.7)	6.74 (45.4)	5.11 (26.1)
5	10.4 (108)	16.0 (256)	20.5 (420)	22.8 (520)	22.9 (524)	21.1 (445)	18.1 (328)
6		27.8 (773)	39.9 (1590)	49.9 (2490)	55.7 (3100)	56.9 (3230)	53.9 (2910)
7		45.5 (2070)	72.8 (5300)	100 (10^4)	123 (151×10^2)	138 (190×10^2)	142 (202×10^2)
8			128 (164×10^2)	191 (365×10^2)	254 (645×10^2)	308 (949×10^2)	345 (119×10^3)
9			214 (458×10^2)	346 (120×10^3)	499 (249×10^3)	647 (419×10^3)	778 (605×10^3)
10			348 (121×10^3)	605 (366×10^5)	931 (867×10^3)	1280 (164×10^4)	1660 (276×10^4)
12					2950 (868×10^4)	4540 (206×10^5)	6620 (438×10^5)
14						144×10^2 (207×10^6)	231×10^2 (534×10^6)

Table V.
$$\varepsilon\{N^2(y;t)\}$$
$\sigma^2(y;t)$ is given in brackets.

t / y	2	3	4	5	6	7	8
2	4.41 (.58)	3.6 (.8)	2.24 (.65)				
3	16.9 (2.6)	20.0 (.2)	17.3 (1.8)	12.4 (1.8)	7.77 (1.77)	4.50 (1.11)	
4	51.4 (9.0)	83.9 (9.4)	99.8 (7.8)	94.0 (6.0)	78.0 (8.3)	51.1 (5.6)	31.8 (5.7)
5	132 (24)	293 (37)	459 (39)	552 (32)	546 (22)	464 (19)	345 (17)
6		890 (117)	1760 (170)	2660 (170)	3200 (100)	3300 (70)	2960 (50)
7		2470 (400)	6000 (700)	109×10^2 (9×10^2)	159×10^2 (8×10^2)	196×10^2 (6×10^2)	205×10^2 (3×10^2)
8			185×10^2 (21×10^2)	403×10^2 (38×10^2)	685×10^2 (40×10^2)	992×10^2 (43×10^2)	121×10^3 (3×10^3)
9			516×10^2 (58×10^2)	133×10^3 (12×10^3)	265×10^3 (16×10^3)	446×10^3 (27×10^3)	631×10^3 (26×10^3)
10			139×10^3 (18×10^3)	417×10^3 (51×10^3)	939×10^3 (72×10^3)	180×10^4 (16×10^4)	290×10^4 (14×10^4)
12					966×10^4 (98×10^4)	233×10^5 (27×10^5)	473×10^5 (35×10^5)
14						234×10^6 (27×10^6)	585×10^6 (51×10^6)

Table VI.

$$\sigma^2/\sigma_F^2$$

σ^2/σ_P^2 is given in brackets.

t / y	2	3	4	5	6	7	8
2	.31 (.30)	.72 (.48)	2.0 (.52)				
3	.25 (.69)	.05 (.05)	.16 (.46)	.25 (.55)	.50 (.72)	.93 (.65)	
4	.24 (1.4)	.14 (1.1)	.095 (.81)	.076 (.64)	.14 (.99)	.14 (.83)	.27 (1.1)
5	.25 (2.3)	.15 (2.3)	.10 (1.9)	.064 (1.4)	.044 (.96)	.044 (.90)	.055 (.94)
6		.16 (4.2)	.11 (4.3)	.070 (3.4)	.032 (1.8)	.022 (1.2)	.017 (.94)
7		.20 (8.8)	.13 (9.6)	.091 (9.0)	.053 (6.5)	.032 (4.3)	.015 (2.1)
8			.13 (16)	.10 (20)	.062 (16)	.045 (14)	.025 (8.7)
9			.13 (27)	.11 (38)	.065 (32)	.065 (42)	.043 (33)
10			.15 (52)	.14 (84)	.083 (77)	.10 (130)	.051 (84)
12					.11 (330)	.13 (590)	.079 (530)
14						.13 (19×10^2)	.096 (22×10^2)

References

1) Alladi Ramakrishnan and P. M. Mathews, Prog. Theor. Phys. (Japan), **11** (1954), 95.
2) Alladi Ramakrishnan, Journal Roy. Stat. Soc. B (London), **8** (1951), 131.

Letters to the Editor

Chain Configurations and Rubber Elasticity

Ei Teramoto

Department of Physics, Kyoto University

June 2, 1954

The Gaussian distribution function of the end-to-end distance of a long chain molecule has two incommodious weakness. That is, the Gaussian distribution function can be applied only to the case in which the end-to-end distance is sufficiently small compared with the whole length of the chain, so it does not express the behavior of large extension. Particularly, the true probability function must vanish at the length of complete extension, nevertheless, the Gaussian function has still a finite value at that point. Secondly, the interactions between the chain elements are not being taken into consideration in this approximation and it is within the range of possibility that if we take into account these interactions the distribution function changes its form essentially from that of the Gaussian function. In order to reform one of these disability, in this letter, we shall discuss a new statistical model of a long chain molecule, in which the distribution function is reduced to the Gaussian function in the region of small extension and also vanishes at the point of complete extension.

Let us consider a long ideal chain composed of N elements of length a, and its two ends are fixed at distance r from each other. Then all the points, which can be reached at least by one of these chain elements, cover the inner space of an ellipsoid whose two principal axes are $X = Na/2$ and $Y = \sqrt{(Na)^2 - r^2}/2$. The volume of this ellipsoid is given by

$$V = 4/8 \cdot \pi X Y^2 = \pi/6 \cdot (Na)^3 \{1 - (r/Na)^2\}. \quad (1)$$

Next we shall introduce another ellipsoid, contracting the above axes uniformly at the ratio $\delta < 1$. The volume of this small ellipsoid, which we shall call the effective ellipsoid, becomes

$$V_{\text{eff}} = \delta^3 V = \pi/6 \cdot (Na\delta)^3 \{1 - (r/Na)^2\}. \quad (2)$$

Now we accept the following assumption, although which seems to be rather rough approximation. The statistical properties of a long chain can be approximated by that of the ideal gas of N particles contained in the above effective ellipsoid, and we assume that the parameter of contraction δ does not depend on r. The configurational partition function of this ideal gas in the balloon is given by

$$\Omega(r) = V_{\text{eff}}^N = (\pi \delta^3/6)^N (Na)^{3N} \{1 - (r/Na)^2\}^N, \quad (3)$$

and the probability that the focal distance of the original ellipsoid is r is given by

$$P(r) = \Omega(r)/\Omega \quad (4)$$

and

$$
\begin{aligned}
\Omega &= \int_0^{Na} \Omega(r) 4\pi r^2 dr \\
&= \left(\frac{\pi \delta^3}{6}\right)^N (Na)^{3(N+1)} 4\pi \int_0^1 (1-x^2)^N x^2 dx \\
&= \left(\frac{\pi \delta^3}{6}\right)^N (Na)^{3(N+1)} 4\pi \cdot \frac{1}{2} B\left(N+1, \frac{3}{2}\right) \\
&= \left(\frac{\pi \delta^3}{6}\right)^N (Na)^{3(N+1)} 4\pi \cdot \frac{2 \cdot 4 \cdot 6 \cdots 2N}{3 \cdot 5 \cdot 7 \cdots (2N+3)},
\end{aligned} \quad (5)
$$

where $B(N+1, 3/2)$ is the Beta function and for sufficiently large value of N we obtain $B(N+1, 3/2) = \sqrt{\pi}/2 \cdot N^{-3/2}$. Hence

$$P(r) = \{2\pi (Na)^3 B(N+1, 3/2)\}^{-1} \{1 - (r/Na)^2\}^N \quad (6)$$

and for the region of small extension, that is, in the case $r \ll Na$, we find

$$P(r) = (1/\pi Na^2)^{3/2} e^{-r^2/Na^2}. \quad (7)$$

Comparing this expression with the usual Gaussian distribution function

$$P_g(r) = (3/2\pi Na^2)^{3/2} e^{-3r^2/2Na^2}, \quad (8)$$

we can find that in the range of small extension the balloon model is reduced to the Gaussian chain whose effective length of segment is $\sqrt{3/2}\, a$. As to the mean square end-to-end distance in our model can be readily evaluated, and we obtain

$$\langle r^2 \rangle = \int_0^{Na} r^2 P(r) 4\pi r^2 dr$$

$$= 4\pi (Na)^5 \{2\pi (Na)^3 B(N+1, \tfrac{3}{2})\}^{-1}$$

$$\times \frac{1}{2} \int_0^1 (1-t)^N t^{3/2} dt$$

$$= \frac{B(N+1, \tfrac{5}{2})}{B(N+1, \tfrac{3}{2})} N^2 a^2$$

$$= \frac{3}{(2N+5)} N^2 a^2 = \frac{3}{2} N a^2 , \qquad (9)$$

which differs again from that of the usual Gaussian chain only by the factor 3/2.

With the use of this distribution function we can obtain the stress-strain relation of the simple elongation of rubber, in just the same manner as the usual treatment of rubber elasticity. The result obtained for the regular cubic network is given by

$$f = A \left\{ \frac{a}{1 - Ba^2} - \frac{1}{a^2(1 - B/a)} \right\} , \qquad (10)$$

where f is the tensile force and a is the ratio of the lengths in the direction of extension after and before extension. And A is a constant which depends upon the structural properties of network and upon the temperature linearly, and another constant $B = (r_0/Na)^2$ where r_0 is the mean length of chains in the natural state. Equation (10) should be compared with that of the Gaussian network
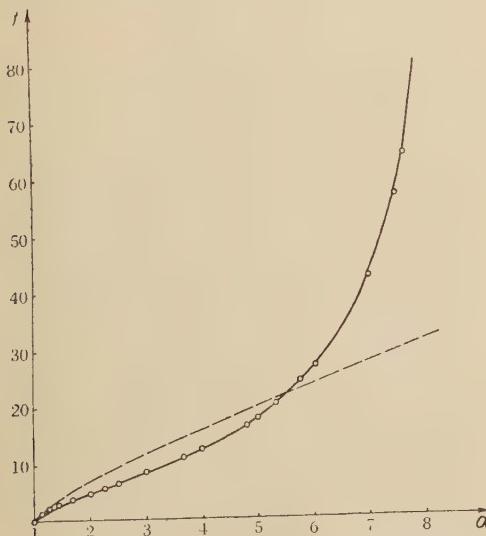
Fig. 1 —— equation (10), ------ Gaussian approximation, ○ experiments (Treloar)

$$f = A(a - 1/a^2). \qquad (11)$$

The stress-strain relation with values $A=2.5$, $B=(1/9)^2$ in equation (10) is plotted in Fig. 1, comparing with the Gaussian approximation and experiments.

On Fermi's Theory of High Energy Nucleon-nucleon Collisions

V. S. Nanda

Department of Physics and Astronomy, Ohio State University

June 5, 1954

At present there is no really satisfactory of multiple production of pions in a high energy nucleon-nucleon collision due to certain inherent difficulties in the meson field theories. Among the various attempts to bypass them the procedures adopted by Lewis et al.[1] and Fermi[2] deserve special attention. Some of the basic assumptions of Fermi's theory have been criticized by Takagi[3], Lewis[4], and more recently by Bhabha[5]. The object of the present note however is to show that, even if Fermi's assumptions are accepted in their entirety, the final results are considerably modified by following a more rigorous thermodynamic procedure.

According to the usual Bose-Einstein theory, N_r the number of particles in the state of energy ϵ_r is given by

$$N_r = \frac{1}{\frac{1}{A} e^{\epsilon_r/kT} - 1} , \qquad (1)$$

where the degeneracy parameter $A=1$ for the pion assembly. In the extreme relativistic case Fermi uses the expression

$$N = \frac{4\pi g}{(Ch)^3} \sim \int_0^\infty \frac{\epsilon^2 d\epsilon}{e^{\epsilon/kT} - 1} , \qquad (2)$$

for the total number of pions, where $\sim$ is the volume in which the energy of the impinging nucleons is concentrated. Its value is given by

$$\sim = (4\pi/3) R^3 \{2Mc^2/W\}, \qquad (3)$$

where $R = \xi \, \hbar/\mu c$, ξ being a dimensionless quantity of the order of unity and μ the pion mass. It might be noted that in obtaining (2) from (1) the summation has been replaced by integration and for the

density of energy levels the asymptotic expression

$$a(\epsilon)d\epsilon = \sim 4\pi g \, \epsilon^2/(Ch)^3 d\epsilon \qquad (4a)$$

has been employed. A more precise result[6], however, gives

$$a(\epsilon)d\epsilon = \left\{ \sim \frac{4\pi g}{(Ch)^3}\epsilon^2 + \theta \, s \frac{\pi g}{2} \frac{\epsilon}{(Ch)^2} + \cdots \right\}, \quad (4b)$$

where the second term is of the nature of a correction, θ being -1 or $+1$ depending on whether the wave function or its derivative vanish on the boundary. If we substitute (4b) in (1) and carry out the integration the contribution of the second term far outweighs the first for $W \gg Mc^2$, indicating that equation (2) and hence all the results obtained from it are invalid for high energies. This necessitates the reconsideration of the whole question.

The final result depends very strongly on the boundary condition which is chosen. For the case of a particle in a box none of the three quantum numbers is permitted to have zero value. We therefore get a very large null point energy and $N \leq 4.3\xi$ always, as has already been pointed out by Auluck and Kothari[7]. Fermi, however, maintains that it is possible for pions even with zero kinetic energy to be formed inside the small volume (see Lewis[4]) in this connection). This is equivalent to using the free boundary condition which permits zero value for the quantum numbers.

For collisions of very high energy the spheroid is contracted to a disc. It can be readily shown the states which are now important are those for which the quantum number corresponding to the contracted dimesion is set equal to zero. This makes the pion gas effectively two-dimensional. We therefore have

$$N = \frac{2\pi g s}{(Ch)^2} \int_0^\infty \frac{\epsilon d\epsilon}{e^{\epsilon/kS}-1} = 6\pi s \left(\frac{kT}{Ch}\right)^2 \zeta(2), \qquad (5)$$

$$W = \frac{2\pi g s}{(Ch)^2} \int_0^\infty \frac{\epsilon^2 d\epsilon}{e^{\epsilon/kT}-1} = 12\pi s \zeta(3) \left(\frac{kT}{Ch}\right)^2 kT, \quad (6)$$

which gives

$$N \sim 3.7 \zeta^{2/3} (W/Mc^2)^{2/3}. \qquad (7)$$

The error involved in replacing summation by integration in (5) and (6) is easily seen to be negligible.

It is interesting to note that Lewis et al[1]. on the basis of their theory obtain

$$N \sim \left(\frac{fT M}{\pi \mu}\right)^{2/3} (W/Mc^2)^{2/3}, \qquad (8)$$

where f is the coupling constant and γ the energy loss of the incident particle. Thus (2/3) power law

for multiplicity results from the two theories inspite of their different starting points. This close resemblance may not be fortuitous. We are examining this point and hope to come back to it in a later publication along with the problem of non-central collisions.

1) H. W. Lewis, J. R. Oppenheimer and S. A. Wouthuysen, Phys. Rev. **73** (1948) 127.
2) E. Fermi, Prog. Theor. Phys. **5** (1950) 570.
3) S. Takagi, Prog. Theor. Phys. **7** (1952), 123.
4) H. W. Lewis, Rev. Mod. Phys. **24** (1952), 241.
5) H. J. Bhabha, Proc. Roy. Soc. A **219** (1953), 293.
6) K. Husimi, Proc. Phys. Math. Soc. Japan, **21** (1939), 759.
7) F. C. Auluck and D. S. Kothari, Phys. Rev. **90** (1953), 1002.

A Remark on the Infinities due to the New Complex Poles of Modified Propagators

Ryôyû Utiyama and Tsutomu Imamura

Department of Physics Osaka University

In this short note we shall show that the new pole of modified propagators pointed out by G. Feldman is not independent on the divergence of perturbation expansion.

Among many attempts[1] to improve on the results of the perturbation method in meson theory, the introduction of Green functions Δ_F' and S_F' (or modified propagators) in place of Δ_F and S_F seems to be most reasonable and promising from the physical point of view. It was pointed out by G. Feldman[1], however, that the introduction of modified propagators gives rise to a new type of infinities which can not be removed by renormalization. These new infinities are due to new complex poles of the modified propagators. This situation was shown by Feldman by considering a simple model of a modified propagator. The outline of his line of reasoning is as follows.

He adopted a simple propagator having the expression

$$Q(k^2) = \sum_{n=0}^{\infty} \left(\frac{-a^2 e^{i\theta}}{k^2 + a^2 - i\epsilon} \right)^n \qquad (1)$$

in momentum space, and discussed convergence of an integral

$$I=\int\int\int\int_{-\infty}^{\infty} Q(k^2)\frac{d^4k}{(k^2+\varkappa^2-i\epsilon')^3} \quad (2)$$

which is considered to correspond to some type of Dyson graphs. In (1) a, θ, and ϵ are some real positive constants. The symbol $\rightarrow$ in (2) means that the integration with respect to k^0 should be carried out along the real axis in a complex k^0-plane.

According to Feldman, (1) can be summed up and put equal to

$$Q(k^2)=\frac{k^2+a^2-i\epsilon}{k^2+a^2-i\epsilon+a^2e^{i\theta}}. \quad (1)'$$

Substituting with $(1)'$ into (2), one encounters new complex poles

$$k^0=\pm\sqrt{\vec{k}^2+a^2-i\epsilon+a^2e^{i\theta}}$$
$$=\pm\{(\vec{k}^2+a^2+a^2\cos\theta)^2+(a^2\sin\theta-\epsilon)^2\}^{1/4}e^{i\phi},$$
$$tg2\phi=\frac{a^2\sin\theta-\epsilon}{\vec{k}^2+a^2(1+\cos\theta)}. \quad (3)$$

Thus I can be transformed into

$$I=\int\int\int_{-\rightarrow}^{\infty} d\vec{k}\int_{\uparrow} dk^0 Q(k^2)\frac{1}{(k^2+\varkappa^2-i\epsilon')^3}+P \quad (4)$$

where the symbol $\uparrow$ means that the k^0-integration should be carried out along the imaginary axis in the k^0-plane. P in (4) stands for a divergent quantity corresponding to the new complex poles (3).

On the contrary, using the ordinary perturbation method, one gets an integral

$$I'=\sum_{n=0}^{\infty}\int_{-\infty}^{\infty} d\vec{k}\int_{\rightarrow} dk^0\left(\frac{-a^2e^{i\theta}}{k^2+a^2-i\epsilon}\right)^n\frac{1}{(k^2+\varkappa^2-i\epsilon')^3} \quad (5)$$

instead of I. Since every integrand of (5) has only ordinary poles, the path of integration with respect to k^0 can be transformed to that along the imaginary axis. Then I' can be written in a compact form as

$$I'=\int_{-\infty}^{\infty} d\vec{k}\int_{\uparrow} dk^0 Q(k^2)\frac{1}{(k^2+\varkappa^2-i\epsilon')^3}<\infty,$$

on account of the fact that after the transformation of the path of integration denominators of the integrand in (5) are positive definite in the limit ϵ, $\epsilon'\rightarrow0$. Thus Feldman gets

$$I=I'+P=\infty,$$

that is, he arrives at the conclusion that although the

expression I' derived by perturbation method is convergent, the other expression I is divergent, thus the new type of divergence is independent on the divergence of perturbation expansion.

From our point of view, however, $Q(k^2)$ defined by (1) is in general not equal to $(1)'$. Only in the case of

$$\epsilon>a^2, \quad (6)$$

$(1)'$ holds for every set of real values of k^μ and in this case the series (1) turns out to be uniformly convergent with respect to k^μ and the argument ϕ of the pole of Q becomes as

$$-\pi/4<\phi<0,$$

that is, the pole of Q becomes an ordinary one. In such a case, of course, one can easily show the relation

$$I=I'=\int_{-\infty}^{\infty} d\vec{k}\int_{\uparrow} dk^0\frac{Q(k^2)}{(k^2+\varkappa^2-i\epsilon')^3}. \quad (7)$$

The appearance of the curious infinity P is essentially due to the fact that the series (1) is not uniformly convergent. In the proof of the equality (7) we have changed the order of integration and summation. As to this point some care should be taken because the domain of integration is infinite, but in our model it is easily seen that this change of order can be permitted.

Although our conclusion exclusively depends on the special form of our model, it is easily conjectured that similar situation will occur in the ordinary meson theory by considering the following fact. If the new type of poles of modified propagators appear in course of calculation and at the same time if the perturbation expansion of modified propagators uniformly converge with respect to momentum variables, we could scarcely understand from what origin these new poles resuls.

The authors wish to express their cordial thanks to Dr. G. Feldman for communicating his work before publication.

1) G. Feldman, "Modified Propagators in Field Theory". This work will be published in near future.

Note on the Contribution of the Single Scattering to the Lateral Structure of Cascade Showers

Jun Nishimura

Department of Physics, Kobe University

Koichi Kamata

Scientific Research Institute

June 9, 1954

It has been pointed out by many authors[1] that the distribution of the shower particles at large distance from the axis can be explained by the large angle Coulomb scattering. This conclusion, however, should be revised if the finite size of the nucleus would be taken into account.

Let us consider a cascade shower initiated by an electron of energy E_0, and the average number of electrons of energy E in this shower at the depth t' is denoted by $\pi(E_0, E, t')dE$. Then the number of particles of energy E scattered in the layer dt' is $\pi(E_0, E, t')dEd\sigma$ where $d\sigma$ is the probability that an electron of energy E is scattered into the solid angle $2\pi \sin \theta d\theta$ in traversing a distance dt'. $d\sigma$ is given approximately for small angle θ by

$$d\sigma = \frac{1}{2 \ln(181\, Z^{-1/3})}\left(\frac{K}{E}\right)^2 \frac{d\theta}{\theta^3} dt', \qquad (1)$$
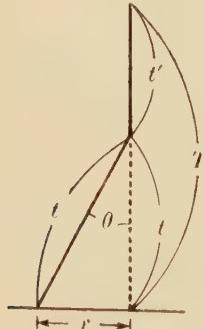
where Z is the atomic number of the traversing material and $K = 21$ Mev. After traversing the additional thickness t, these particles are displaced from the axis by $r = t \sin \theta$, and produce showers with a total number of particles $\Pi(E, 0, t)$ (see Fig. 1). The number of particles in the angular ring between r and $r+dr$ at the depth T is then given by the integral[2] for the point charge model of the nucleus

$$\frac{dr}{r^3}\int_0^{E_0}\int_0^T \pi(E_0, E, t')\frac{1}{2\ln(181\, Z^{-1/3})}\left(\frac{K}{E}\right)^2$$
$$\times \Pi(E, 0, T-t')(T-t')^2 dEdt', \qquad (2)$$

where r, t', T and t are measured in radiation units.

If the charge of the scattering nucleus is uniformly distributed over a volume of the sphere of radius R, the maximum scattering angle θ_m can be represented approximately by[3]

$$\theta_m \simeq \lambda/R \qquad (3)$$

where λ is the de Broglie wave length of the incident electron. The formula (3) turns out to be

$$\theta_m = \frac{\varphi}{E}, \qquad (4)$$

where we put $\hbar c/R = \varphi$ and E is the energy of the incident electron. As $r \leq t \sin \theta_m \leq t\theta_m$, we get

$$E \leq t\varphi/r, \qquad (5)$$

and the integral with respect to energy in (2) vanishes for the energy region $E > \varphi t/r$. Since

$$\Pi(E, 0, t) \simeq 0 \quad \text{for} \quad E \leq \varepsilon t, \qquad (6)$$

(where ε is the critical energy of the traversing material), it follows from the relations (5) and (6),

$$\text{integral (4)} \simeq 0 \quad \text{for} \quad r \geq \varphi/\varepsilon. \qquad (7)$$

Thus the contribution of single scattering to the lateral displacements of the shower particles is small at the distance from the axis larger than φ/ε radiation units. Assuming $R = 1.4 \times 10^{-13}A^{1/3}$cm where A is the atomic weight, φ for air turns out to be about 70 Mev, and $\varphi/\varepsilon \simeq 3/4$ radiation units.

Following the results obtained by Molière[1], the effect of the single scattering has little importance on the structure function within 3/4 radiation units while the behavior of the structure function at the large distance from the axis is mainly determined by the single scattering in the case of air showers. It might be then said that if the cut off angle for the single scattering cross section were "sharp cut off" as assumed in this note, the single scattering would have practically no contribution to the structure function. However, the following points should be put in mind concerning this conclusion:

(1) The scattering cross section for large angles has not "sharp cut off" edge. However, it is sure that, compared with the point charge theory, scattering cross section falls off remarkably outside the "sharp cut off" angle.[4],[5]. Thus even though it seems that our approximation is crude, the results thus

obtained is not so far from the reality.

(2) Plural scattering should be taken into consideration. It is very difficult to treat this problem accurately. In order to test our approximation using the structure function (due to multiple scattering only) instead of the cascade function Π in eq. (4), the result is not subject to important change because more than 60% of the shower particles are concentrated within 0.5 $r_1^{6)}$. The definite conclusion on this point however still remains to be reached.

At any rate it follows from the above discussions that the distribution of air shower particles at large distance from the core should not be interpreted in terms of the contribution of only the single scattering[7], but in another way, for example, in terms of the angular divergence of the parent meson production.

1) For example, G. Molière, *Cosmic Radiation*, edited by W. Heisenberg. (Dover Publications, New York, 1946, Chap. 3) ; L. Eyges, Phys. Rev. **74** (1948), 1801.

2) L. Eyges, See ref. (1).

3) E. J. Williams, Proc. Roy. Soc. **169** (1939), 531.

4) E. M. Lyman, A. O. Hanson and M. B. Scott, Phys. Rev. **84** (1951), 626.

5) R. Hofstadter, H. R. Fechter and J. A. McIntyre, Phys. Rev. **92** (1952) 978.

6) J. Nishimura and K. Kamata, Prog. Theor. Phys. **6** (1951) 628 r_1 is the characteristic scattering distance which is defined as E_s/ε times radiation unit.

7) G. Cocconi ; Phys. Rev. **12** (1947) 350. See also ref. (1).

Some Generalization of the Theory of Scattering with Application to Slow Electrons in Polar Crystals

Akira Morita

Physical Institute, Faculty of Science, Tohoku University

June 10, 1954

The general theory of scattering has been formulated by many authors[1] during recent years. However it is not yet of sufficient generality to cover most cases of interest. It is applicable to the cases, in which the damping factor, Γ_j, mentioned below, (more exactly speaking, $\Gamma_j T$) is infinitesimal, but is inapplicable to the electron-lattice vibration system, in which Γ_j is finite. Therefore we shall reformulate the theory of scattering so as to be applicable also to such a system.

Let us suppose that the eigenvalues of the total Hamiltonian

$$H = H_0 + H', \qquad (1)$$

are E_n and those of the " unperturbed part " H_0 of of the Hamiltonian are ε_n, that is

$$(H_0 + H')\psi_n = E_n \psi_n, \qquad (2)$$
$$H_0 \phi_n = \varepsilon_n \phi_n. \qquad (2)$$

Here we assume that there are no bound states in either case. The energy shift due to the interaction part H', i.e., the self-energy is given by

$$\Delta_n = E_n - \varepsilon_n. \qquad (4)$$

In order to describe the scattering process correctly, H_0 must be modified formally so as to make its energy spectrum coincide with that of $(H_0 + H')$. Let us define an operator Δ, according to Gell-Mann and Goldberger, by

$$\Delta = \sum \phi_n > \Delta_n < \phi_n. \qquad (5)$$

and write

$$H = (H_0 + \Delta) + (H' - \Delta). \qquad (6)$$

Then the new interaction $(H' - \Delta)$ fulfils our purpose, since it produces no energy shift.

As is well known, the solution of scattering problem corresponding to the incident wave

$$\Psi_j = e^{-iE_j t}\phi_j \qquad (7)$$

is given by

$$\Psi_j(t) = e^{-iHt}\Psi_j(0), \qquad (8)$$

where

$$\Psi_j(0) = \phi_j + \frac{1}{E_j - (H_0 + \Delta) + i\epsilon}(H' - \Delta)\Psi_j(0).$$

Now if we suppose that the interaction was switched in at the sufficiently distant past, i.e., at $t = -T$ $(= -\infty)$, we may put

$$\langle \phi_j | \Psi_j(0) \rangle = e^{-1/2 \cdot \Gamma_j(\epsilon)T}. \qquad (10)$$

Here

$$\Gamma_j = \lim_{\epsilon \to 0} \Gamma_j(\epsilon) \qquad (11)$$

is the damping factor of ϕ_j-state. On the strength of eq. (10), we know that the self-energy Δ_j is determined by

$$\Delta_j = \langle \phi_j | (H' - \Delta_j) \frac{P}{E_j - H}(H' - \Delta_j) | \phi_j \rangle \qquad (12)$$

and the relation

$$R_{jj} = \lim_{\epsilon \to 0} \langle \phi_j | H' - \Delta | \Psi_j(0) \rangle$$

$$= \lim_{\epsilon \to 0} \left\{ \langle \phi_j | (H' - \Delta) \frac{1}{E_j - H + i\epsilon} (H' - \Delta)' | \phi_j \rangle - \Delta_j \right\}$$

$$= 0 \qquad (13)$$

holds. Furthermore, let us put

$$\Psi_j(0) = e^{-1/2\Gamma_j T} \Psi_j'(0) \qquad (14)$$

and

$$\bar{R}_{ij} = \langle \phi_i | H' - \Delta | \Psi_j'(0) \rangle, \qquad (15)$$

then we find that the transition probability from the state ϕ_j to the state ϕ_i is given by

$$\frac{2\pi}{h} \delta(E_j - E_i) |\bar{R}_{ij}|^2. \qquad (16)$$

The damping factor Γ_j is given by

$$\Gamma_j = \frac{2\pi}{h} \sum_{i \neq j} \delta(E_j - E_i) |\bar{R}_{ij}|^2 \qquad (17)$$

and $\bar{R}_{ij}$ is determined by the equation

$$\bar{R}_{ij} = B_{ij} + \sum_{k \neq j} \frac{B_{ik} \bar{R}_{kj}}{E_j - E_k + i\epsilon} - \frac{\Delta_i \bar{R}_{ij}}{E_j - E_i + i\epsilon},$$
$$(i \neq j) \qquad (18)$$

where

$$B_{ij} = \langle \phi_i | H' | \phi_j \rangle.$$

Now it must be noted that eq. (18) does not contain any process in which the system returns to the initial state on the way. The meanings of the third term of the right hand side of eq. (18) are as follows. Let us iterate eq. (18) once, then we obtain, as a part of the second term,

$$\left(\sum_{k \neq j} \frac{B_{ik} B_{ki}}{(E_j - E_k + i\epsilon)} \right) \frac{\bar{R}_{ij}}{E_j - E_i + i\epsilon}. \qquad (19)$$

Provided that we continue successive iterations, we can collect similar terms in the form

$$\frac{\Delta_i' \bar{R}_{ij}}{E_j - E_i + i\epsilon} \qquad (20)$$

where

$$\Delta_i' = \sum_{\substack{k \neq j \\ \neq i}} \frac{\bar{B}_{ik} \bar{B}_{ki}}{E_j - E_k + i\epsilon}$$

$$+ \sum_{\substack{k \neq j \\ \neq i}} \sum_{\substack{l \neq j \\ \neq i}} \frac{\bar{B}_{ik} \bar{B}_{kl} \bar{B}_{li}}{(E_j - E_k + i\epsilon)(E_j - E_l + i\epsilon)} + \cdots,$$
$$(21)$$

$$\bar{B}_{ik} = \langle \phi_i | H' - \Delta | \phi_k \rangle. \qquad (21')$$
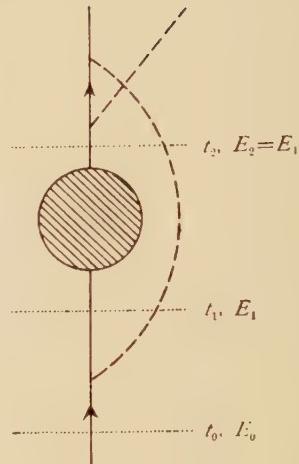
By combining this term with the third term of eq. (18), we obtain

$$\frac{(\Delta_i' - \Delta_i) \bar{R}_{ij}}{E_j - E_i + i\epsilon}. \qquad (22)$$

This quantity becomes zero when $E_i = E_j$, since $\Delta_i' = \Delta_i$ in this case. This fact means that the difficulty of self-energy type does not appear, since we have determined the self-energy Δ_j so that the new interaction part $(H' - \Delta)$ produces no energy shift. For example, the matrix element of $\bar{R}$ corresponding to Fig. 1. has the factor

$$\sum_{k_1} \frac{(\Delta_1' - \Delta_0)}{(E_0 - E_1 + i\epsilon)^2}$$

between time t_1 and time t_2, as easily seen from (18). This factor *does not* contain any singularity, because $\Delta_1' = \Delta_1$ when $E_1 = E_0$. But, if we ignore the third term in eq. (18), the corresponding factor might contain the difficulty of infinity when $E_1 = E_0$.

Now it is clear how to treat the scattering problem, in which there are self-energies and finite damping factors. Our formulations are applicable to the electron-phonon system.

In their papers concerning the mobility of slow electrons in polar crystals, Fröhlich, Mott[2] and Callen[3] calculated, considering only direct processes and using the usual perturbation theoretic methods, scattering probabilities of slow electrons due to processes accompanying absorption or emission of one phonon. Their results are seemed to be in fairly well agreement with experimental data.[4] But, in these crystals, the interaction of conduction electrons with lattice vibrations is so strong that the usual perturbation theoretic methods are inapplicable. In

order to obtain any quantitative information on this problem, we must take higher order processes besides direct processes into consideration.

We applied our formulation to this problem with some approximation concerning the number of virtually emitted phonons and the assumption that temperatures are so low that the possibility of absorption of real phonon is negligibly small, and obtained the following conclusions.

Although the interaction is so strong that the perturbation theoretic methods are inapplicable, the results obtained by Fröhlich and others by using the first order approximation of the perturbation theory are, if the effective mass is adopted in the place of the electron mass, almost equal to the results obtained by the more rigorous present method at least in the cases where electron energies are very small compared with energy quantum of phonon.

A detailed account of this work will be publisched in the Science Reports of Tohoku University, Series I.

1) See, for example, B. Lippmann and J. Schwinger, Phys. Rev. **79** (1950), 467 or M. Gell-Mann and M. L. Goldberger, Phys. Rev. **91** (1953), 398.

2) H. Fröhlich and N. F. Mott, Proc. Roy. Soc. **171** (1939), 496.

3) H. Callen, Phys. Rev. **76** (1949), 1394.

4) A. G. Redfield, Phys. Rev. **91** (1953), 753.

Pion-nucleon Scattering by Tamm-Dancoff Method

Kin-ichi Ishida and Atsushi Takahashi

Physical Institute, Faculty of Science, Tohoku University

June 11, 1954

The pion-nucleon scattering for the isotopic spin $I=3/2$ state in Tamm-Dancoff approximation has been treated by Fubini[1], and the phase shifts thus obtained considerably agreed with the experiment[2], for the coupling constant $g^2/4\pi=10.5$ and the cut-off momentum$=0.92M$, the cutting off being assumed in order to take account of the relativistic effects such as recoil non-relativistically. However, similar treatmentment for $I=1/2$ state has not been available so far[3], since the mass and charge renormalization is necessary. We have done this keeping correspondence with the renormalization procedure in covariant formalism[4], and calculated the phase shifts

for $I=1/2$ state in the same approximation as Fubini's, or taking into account all graphs with $n \leqq 1$ nucleon pairs and $m \leqq 2$ mesons except for graphs with closed loops, the graphs involving meson nucleon interactions in the state with the nucleon pairs and the graphs in Fig. 1.

We first eliminate the nucleon pair states from the equations of motion, and discuss the problem comparing with the Dyson's and Foldy's transformation. The interaction Hamiltonian consists of three parts: scattering H_s, pair creation H_c and pair annihilation H_a of nucleons. The integral equation for the state vector ψ with the nucleon pair eliminated, turns out

$$(W-H_0+\varDelta)\psi = H_s\psi + H_a \frac{1}{W-H_0} H_c\psi. \quad (1)$$

The self energy term $\varDelta$ is to be cancelled out. The last term contains scattering, pair creation and pair annihilation of mesons, as well as the self energy effects. If the nucleon recoil is completely neglected, this term gives the well-known core term[5]

$$g^2/2M \int \psi^* \beta \phi^2 \psi dv. \quad (2)$$

If we go up to the order k/M, there appears, beside the p-wave interaction, the τ-spin dependent term derived by Drell and Henley[5]:

$$H_{\mathrm{II}} \equiv (g/2M)^2 \int \psi^* \pi [\tau \times \phi] \psi dv. \quad (3)$$

This arises from the meson energy dependence in the denominators of the intermediate states involving nucleon pairs. H_{II} separates the phase shifts a_1, a_3; this separation is not large enough to change the sign of a_1, so far as the cut-off procedure is assumed.

H_a, H_c and H_s are divided into creation H^+ and annihilation operator H^- of mesons. Then, in the C. M. system we get the following equations for the state ψ_0 with no meson and the state ψ_1 with one meson:

$$\left[W-M+\varDelta_0 - {}_0\left(H_a \frac{1}{W-H_0} H_c^+ \right)_0 \right]\psi_0$$
$$= \sum \langle 0|H_s|\vec{k}\rangle\langle\vec{k}|\psi_1\rangle, \quad (4a)$$

$$\left[W-E_k-w_k+\varDelta_1 - {}_0\left(H_s \frac{1}{W-H_0} H_s^+ \right)_0 \right.$$
$$\left. - {}_0\left(H_a \frac{1}{W-H_0} H_c^+ \right)_0 \right]\langle\vec{k}|\psi_1\rangle$$
$$= \langle\vec{k}|H_s^+|\psi_0\rangle + \sum\langle\vec{k}|\left(H_a \frac{1}{W-H_0} H_c^+ \right)$$

$$+ H_a{}^+ \frac{1}{W - H_0} H_c{}^- + \left(H_s - \frac{1}{W - H_0} H_s{}^+ \right)' |\vec{k}'\rangle$$
$$\cdot \langle \vec{k}' | \psi_1 \rangle, \qquad (4b)$$

where $_0(\ )_0$ represents the self energy effects. The bracket on the left side in (4b) is to be renormalized so as to allow the existence of the freely propagating meson-nucleon system. Then Δ_1 is determined, by the condition that for $W = E_k + w_k$ the bracket is zero, thus having the form: $[W - E_k - w_k]$ $\cdot [1 + P(W - E_k - w_k)]$. P is different from zero for $W = E_k + w_k$, and this $P(0)$ is to be charge-renormalized. We note that $P(0)$ is just the meson-dissociation probability of the nucleon, and stands with the renormalized coupling constant g_1 as follows:

$$g_1 = g/(1 + P(0)). \qquad (5)$$

If we define the modified propagator $1/a'$, where
$$a' = [W - E_k - w_k] [1 + P(W - E_k - w_k) - P(0)], \qquad (6)$$
the integral equation for $\langle \vec{k} | \psi_1 \rangle$ is written as follows:

$$\langle \vec{k} | \psi_1 \rangle = \delta(\vec{k} - \vec{k}_0) + \frac{1}{a'} \langle \vec{k} | H_s{}^+ | \psi_0 \rangle +$$
$$+ \frac{1}{a'} \sum \langle k | \left(H_a - \frac{1}{W - H_0} H_c{}^{+\prime} \right)' + H_a{}^+ \frac{1}{W - H_0} H_c{}^-$$
$$+ \left(H_s - \frac{1}{W - H_0} H_s{}^+ \right)' |\vec{k}'\rangle \langle k' | \psi_1 \rangle. \qquad (7)$$

The solution of this equations may be written in the form

$$\langle \vec{k} | \psi_1 \rangle = \langle \vec{k} | K | \vec{k}_0 \rangle + \langle \vec{k} | K \frac{1}{a'} H_s{}^+ | \psi_0 \rangle, \qquad (8)$$

where:
$$\langle \vec{k} | K | \vec{k}' \rangle = \delta(\vec{k} - \vec{k}')$$
$$+ \frac{1}{a'} \sum \langle \vec{k} | \left(H_a - \frac{1}{W - H_0} H_c{}^+ \right)' + H_a{}^+ \frac{1}{W - H_0} H_c{}^-$$
$$+ \left(H_s - \frac{1}{W - H_0} H_s{}^+ \right)' |\vec{k}''\rangle \langle \vec{k}'' | K | \vec{k}' \rangle. \qquad (9)$$

In (8), the last term in the right represents a process in which one meson is emitted by a nucleon, so that this term is renormalized by the condition: if the freely propagating nucleon with vanishing momentum interacts with the external meson with vanishingly small momentum and energy, the vertex acts as the bare one. Substituting $\langle \vec{k} | \psi_1 \rangle$ thus obtained to (4a), we shall get an equation consisting of two terms; the one involving the state vector ψ_0 and the other with the incident wave $|\vec{k}_0\rangle$. Now, for the latter which represents absorption of the incident meson, we can apply the renormalization for vertex as before,

while for the former we do the renormalization for nucleon propagator as for $\langle \vec{k} | \psi_1 \rangle[6]$. Inserting ψ_0 thus obtained, we shall get the required $\langle \vec{k} | \psi_1 \rangle$.

Employing this solution, we calculated the phase shifts for $I = 1/2$ state and 235 Mev pion energy,

in the 1st approximation in Fredholm sense and adopting the coupling constant and the cut-off momentum of Fubini:

	a_1	a_{11}	a_{13}
Theory	-20^0	-6.5^0	-4.7^0
Experiment	$+10^0$	-5^0	$+2^0$.

The contribution from $H_{\rm II}$ is small, and therefore neglected entirely. This implies the result $a_1 = a_3$, in disagreement with the experimental one.

It seems to us that this disagreement is due to the following fact. Since the the s-wave interaction is only possible through the intermediate states involving nucleon pair, the range of s-wave interaction is very short ($\lesssim \hbar/2Mc$). Thus nucleons are in very short distances, and may interact strongly. This has been also showed in the calculation of nuclear force by Brueckner and Watson[7].

The result of calculation shows that the vertex correction in the Tamm-Dancoff approximation is considerably small, even though we take the last two graphs in Fig. 1. The amount of the charge to be renormalized in the nucleon propagator as well as in vertex is about 10% of the bare one, showing that the charge renormalization is practically needless. However, we must remark that the smallness of $P(0)$ is due to the cancellation of the two contributions from the state of one nucleon and one meson and the state of one nucleon, one nucleon pair and one meson.

The authors wish to thank Professor K. Nakabayasi for encouragement and advice and also Mr. S. Inawashiro and Mr. K. Yomogita for valuable discussions.

1) S. Fubini, Nuovo Cimento **10** (1953), 564.
2) H. L. Anderson, E. Fermi, R. Martin and D. E. Nagle, Phys. Rev. **91** (1953), 155.
3) This problem has been recently treated in slightly different way by Chiba et al. independently of

ours. Cf. S. Chiba, M. Yamazaki and N. Fukuda, Soryushiron Kenkyu (Mimeographed circular in Japanese) **6** (1954), 753.

4) S. Fubini, Nuovo Cimento **10** (1953), 851; K. Sawada, S. Okubo and S. Matsuyama, Soryushiron kenkyu (Mimeographed circular in Japanese) **5** (1953), 844.

5) S. D. Drell and F. Henley, Phys. Rev. **88** (1952), 1053.

6) The treatment of overlapping divergence performed by Chiba et al., l.c., is in our case not clear as well as in Fubini, though its effect seems to be comparatively small at least in the cut-off theory. This point requires closer studies, being now in progress.

7) K. A. Brueckner and K. M. Watson, Phys. Rev. **92** (1953), 1023.

The General Theory of Relativity and the Expanding Universe

Shambhunath Bhattacharya

Bhangabasi College, Calcutta, India

June 18, 1954

In this short note I have shown that Einstein's spacetime is a particular aspect of a conformal Riemannian spacetime and consequently Einstein's gravitational field equation of a particle which so superbly explains local gravitational phenomena is a special part of the general conformal field equation. So quite naturally this theory supports Einstein's predictions. Furthermore the possibility of an expanding universe has shown in the same spacetime by means of a simple linear differential equation.

$\overline{ds^2} = e^{2\sigma} g_{ij} dx_i dx_j = e^{2\sigma} ds^{2*}$ is the metric of the conformal spacetime where σ is a scalar function of coordinates. It is always possible to write the metric as

$$\overline{ds^2} = e^{2\sigma} (c^2 dt^2 - dx_1^2 - dx_2^2 - dx_3^2) \tag{1}$$

in terms of a suitable coordinate system, where c is

* Eisenhart's Riemannian Geometry. p. 89 (Second printing, 1949)

the constant velocity of light and t the time. Consequently equation of geodesic in this new spacetime may be written as

$$\frac{d}{ds}\left(e^{\sigma}\frac{dx_i}{ds}\right) = g^{lk} e^{\sigma},_k \tag{2}$$

where $\partial e^{\sigma}/\partial x_k = e^{\sigma},_k$.

Let a perfect, homogeneous fluid uniformly spread through the universe. We consider the universe as a whole and write the field-equation as

$$\overline{R} = 6e^{-2\sigma}(\Delta_2\sigma + \Delta_1\sigma) = \varkappa\overline{T} = \varkappa e^{-2\sigma}\rho = 6P^{-3}\Delta_2 P \tag{3}*$$

where $P = e^{\sigma}$, $\overline{T}$ is the invariant of the energy-momentum tensor, $\overline{R}$ the curvature invariant.

The energy-momentum tensor of the fluid after Einstein** is

$$\overline{T}_{ij} = -\overline{g}_{ij} p + \overline{g}_{i\alpha} g_{i\beta} \ dx_\lambda/d\overline{s} \cdot dx_\beta/d\overline{s} \cdot \rho$$

where p the pressure and ρ the density of the field producing fluid. We assume that pressure is negligible compared to the density of the fluid. Then $\overline{g}_{ij}\overline{T}_{ij} = \overline{T} = e^{-2\sigma}\rho$. From the considerations of the dimension of a physical quantity we can show that $e^{\sigma}(=P)$ has the dimension L^{-1} where L denotes length. A non-singular solution of the equation $\Delta_2 P = (\varkappa\rho P)/6$ depends only on time and is

$$e^{\sigma} = P = Ae^{\alpha t} + Be^{-\alpha t}, \text{ where } \alpha = \sqrt{\varkappa\rho/6}.$$

If $\rho = 0$ we get space-time corresponding to $R^l{}_{ijk} = 0$ and for the field on an isolated particle we can wish the Ricci tensor $R_{ij} = 0$ because density of matter is negligibly small there and e^{σ} is very nearly equal to 1.

If $e^{\sigma} = e^{-\alpha t}$ and $v = $ radial velocity $= dr/dt = c \tanh \phi$ where ϕ is a new parameter with the help of (2) and we see that $v = c \tanh(dr/c) \approx \alpha r$, the velocity distance equation if higher powers of $1/c$ are neglected. The limit velocity is c.

The metric of the universe is $e^{-2\alpha t}(c^2 dt^2 - dx_1^2 - dx_2^2 - dx_3^2)$. The universe has an origin of time, the starting point of expansion. The cylindrical universe of Einstein might have existed just before starting to expand.

* Eisenhart's *Riemannian Geometry*. p. 90 for $\overline{R}$ and p. 41 for $\Delta_2\sigma$, $\Delta_1\sigma$ etc.

** *The Principle of Relativity* (Dover), p. 152. also Tolman's *Relativity, Thermodynamics and Cosmology* (Oxford).

INDEX TO VOLUME 11